

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 08:24:10 ; Search time 45.7 Seconds
(without alignments)
54.389 Million cell updates/sec

Title: US-09-252-828A-2

Sequence: 1 SWFPVQGPADICQCCNKGDC.....TPSHSRQPHVMSQMSRSVS 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_0601: *
2: /SIDSB/gcgdata/geneseq/geneseqp/AA1980.DAT: *
3: /SIDSB/gcgdata/geneseq/geneseqp/AA1981.DAT: *
4: /SIDSB/gcgdata/geneseq/geneseqp/AA1982.DAT: *
5: /SIDSB/gcgdata/geneseq/geneseqp/AA1983.DAT: *
6: /SIDSB/gcgdata/geneseq/geneseqp/AA1984.DAT: *
7: /SIDSB/gcgdata/geneseq/geneseqp/AA1985.DAT: *
8: /SIDSB/gcgdata/geneseq/geneseqp/AA1986.DAT: *
9: /SIDSB/gcgdata/geneseq/geneseqp/AA1987.DAT: *
10: /SIDSB/gcgdata/geneseq/geneseqp/AA1988.DAT: *
11: /SIDSB/gcgdata/geneseq/geneseqp/AA1989.DAT: *
12: /SIDSB/gcgdata/geneseq/geneseqp/AA1990.DAT: *
13: /SIDSB/gcgdata/geneseq/geneseqp/AA1991.DAT: *
14: /SIDSB/gcgdata/geneseq/geneseqp/AA1992.DAT: *
15: /SIDSB/gcgdata/geneseq/geneseqp/AA1993.DAT: *
16: /SIDSB/gcgdata/geneseq/geneseqp/AA1994.DAT: *
17: /SIDSB/gcgdata/geneseq/geneseqp/AA1995.DAT: *
18: /SIDSB/gcgdata/geneseq/geneseqp/AA1996.DAT: *
19: /SIDSB/gcgdata/geneseq/geneseqp/AA1997.DAT: *
20: /SIDSB/gcgdata/geneseq/geneseqp/AA1998.DAT: *
21: /SIDSB/gcgdata/geneseq/geneseqp/AA1999.DAT: *
22: /SIDSB/gcgdata/geneseq/geneseqp/AA2000.DAT: *
23: /SIDSB/gcgdata/geneseq/geneseqp/AA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	242	100.0	41	20	AAV30214
2	235	97.1	424	21	AAV52988
3	229	94.6	372	21	AAV67287
4	226	93.4	223	15	AAV65552
5	226	93.4	223	20	AAV42483
6	226	93.4	223	20	AAV81820
7	226	93.4	223	21	AAV82218
8	226	93.4	223	21	AAV52692
9	226	93.4	223	21	AAV52183
10	226	93.4	223	21	AAV52987
11	225	93.0	372	13	AAV22239

12	192	79.3	424	15	AAV53498
13	151	62.4	426	15	AAV55199
14	151	62.4	426	20	AAV42472
15	151	62.4	426	20	AAV81809
16	151	62.4	426	21	AAV82207
17	151	62.4	426	21	AAV52681
18	151	62.4	426	21	AAV52172
19	151	62.4	426	21	AAV52976
20	140	57.9	426	15	AAV48068
21	137	56.6	424	15	AAV60165
22	137	56.6	424	15	AAV47198
23	136	56.2	424	15	AAV55202
24	136	56.2	424	20	AAV42475
25	136	56.2	424	20	AAV81812
26	136	56.2	424	21	AAV82210
27	136	56.2	424	21	AAV52684
28	136	56.2	424	21	AAV52175
29	136	56.2	424	21	AAV52979
30	126	52.1	424	11	AAV07058
31	119	49.2	258	14	AAV41004
32	119	49.2	258	15	AAV69551
33	119	49.2	420	15	AAV69550
34	119	49.2	421	15	AAV55196
35	119	49.2	421	20	AAV42469
36	119	49.2	421	20	AAV81806
37	119	49.2	421	21	AAV82204
38	119	49.2	421	21	AAV52678
39	119	49.2	421	21	AAV52169
40	119	49.2	421	21	AAV52973
41	109	45.0	421	15	AAV55205
42	109	45.0	421	20	AAV42478
43	109	45.0	421	20	AAV81815
44	109	45.0	421	21	AAV82213
45	109	45.0	421	21	AAV52687

ALIGNMENTS

RESULT 1
ID AAV30214 standard; peptide: 41 AA.
XX AAV30214;
XX 01-NOV-1999 (first entry)
XX
XX Amino acids 308 to 348 of human zona pellucida protein 3 (hzp3).
XX Human; zona pellucida protein 3; hzp3; glycoprotein; human spermatozoa;
XX acrosome reaction; fertility; sperm motility; fertilization.
XX
XX Homo sapiens.
XX
XX PN M09942581-A1.
XX PD 26-AUG-1999.
XX PF 19-FEB-1999; 99WO-US03273.
XX PR 19-FEB-1998; 98US-0075079.
XX
XX (EVIR-) EASTERN VIRGINIA MEDICAL SCHOOL.
XX PA Dong K, Gibbons WE, Oehninger S;
XX PI WPI; 1999-527476/44.
XX DR New recombinant human zona pellucida protein 3 - used to develop
XX PT products for diagnostic and therapeutic use related to human
XX PS fertility
PS Claim 7; Fig 1; 35pp; English.

XX The present sequence represents amino acids 308 to 348 of human zona
CC pellucida protein 3 (hZP3). The protein is a glycoprotein, and can the
CC strongly bind human spermatozoa and induce an acrosome reaction in the
CC spermatozoa. The specification describes methods for producing
CC recombinant hZP3. The protein can provide high levels of biologically
CC active hZP3 of consistent quality for use in clinical and research
CC chemistry, e.g. diagnostic and therapeutic use related to human
CC fertility. The protein retains the ability to bind spermatozoa and induce
CC the acrosome reaction. The hZP3 can be used to treat sperm to allow the
CC sperm to respond more efficiently to motility and acrosome reaction
CC regulators within the female reproductive tract and increase the rate
CC of successful fertilization.

SO Sequence 41 AA;

Query Match 100.0%; Score 242; DB 20; Length 41;
Best Local Similarity 100.0%; Pred. No. 8.6e-23;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SWPVGPDADICCCNKGDCGTPSHSRPHVMQWSRSYS 41
1 SWIPVGPADICGCKNGYDCGTPSHSRPHVMQWSRSYS 41
|||||

RESULT 2
AAV52988
ID AAV52988 standard; Protein: 424 AA.
AC AAV52988;
DT 14-FEB-2000 (first entry)
XX
DE Human zona pellucida ZPC protein sequence.
KW Zona pellucida; ZPA; ZPB; ZPC; Infertility; sterility; vaccine;
KW Immunoc contraceptive; contraception; immunisation.
OS Homo sapiens.
XX
PN US5976545-A.
XX
PD 02-NOV-1999.
XX
PF 07-JUN-1995; 95US-0484158.
XX
PR 09-NOV-1993; 93US-0149223.
PR 09-NOV-1992; 92US-0973341.
PR 29-JAN-1993; 93US-0012990.
XX
PA (ZONA-) ZONAGEN INC.
XX
PI Podolski JS, Hsu KT, Harris JD;
XX
DR WPI; 2000-037080/03.
DR N-PSDB; AA233285.
XX
PT New recombinant mammalian zona pellucida C proteins, for induction of
PT female sterility -
XX
PS Disclosure; Column 153-156; 87pp; English.
XX
CC The present invention describes recombinant zona pellucida (ZP)
CC proteins, specifically ZPC proteins. Also described are fusion proteins
CC of ZPC with any of keyhole limpet haemocyanin, muramyl dipeptide,
CC histidine tag, beta-galactosidase or palmitic acid, capable of
CC stimulating mammalian antibodies that recognise mammalian ZPC.
CC Recombinant ZPC proteins are used to immunise animals to induce transient
CC infertility or permanent female sterility, in humans or other mammals.
CC AA233243 to AA233254, AA233270 to AA233274 and AA233285 encode mammalian
CC ZP proteins given in AAV52988 to AAV52988 from the present invention.
CC AA233255 to AA233269 and AA233275 to AA233284 represent oligonucleotides

CC used in the exemplification of the present invention.

SO Sequence 424 AA;

Query Match 97.1%; Score 235; DB 21; Length 424;
Best Local Similarity 95.1%; Pred. No. 5.9e-21;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWPVGPDADICCCNKGDCGTPSHSRPHVMQWSRSYS 41
1 SWIPVGPADICGCKNGYDCGTPSHSRPHVMQWSRSYS 348
|||||

RESULT 3
AAV67287
ID AAV67287 standard; Protein: 372 AA.
AC AAV67287;
DT 05-APR-2000 (first entry)
XX
DE Human zona pellucida protein (ZP3) amino acid sequence.
XX
KW Zona pellucida protein; ZP3; sperm receptor; fertilisation; antiserum;
KW sperm receptor; antigenic molecule; sperm binding factor; vaccine;
KW sterilisation; fertility modulation.
XX
OS Homo sapiens.
XX
PN WO965520-A1.
XX
PD 23-DEC-1999.
XX
PF 18-JUN-1999; 99WO-US13858.
XX
PR 19-JUN-1998; 98US-0089950.
XX
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
XX
PI Herr JC, Coonrod SA;
XX
DR WPI; 2000-106016/09.
DR N-PSDB; AA257790.
XX
PT Antigenic molecule used in methods for modulating fertility -
XX
PS Claim 6; Fig 12B; 103pp; English.
XX
CC This is the amino acid sequence of the human zona pellucida protein
CC (ZP3). ZP3 takes part in the primary binding between egg and sperm,
CC during fertilisation. ZP3 binds to receptors on the sperm. The ZP3
CC protein is used in the invention, which relates to an antigenic molecule,
CC which consists of an isolated egg surface protein (e.g. ZP3) covalently
CC linked to a glycosylated phosphatidylinositol moiety and having sperm
CC binding activity. Antiserum comprising an antibody to an egg surface
CC protein, can be prepared by recovering serum with the ability to bind
CC antigen from an animal following immunisation with an immunogenic
CC fragment of an egg surface protein. The antigenic molecule the antiserum,
CC and an antibody against an egg surface protein can be used in methods for
CC modulating fertility. The antibodies can be used in vaccination against
CC fertilisation, sterility, diagnostic immunoassays, passive
CC immunotherapy and generation of antidiotypic antibodies.

SO Sequence 372 AA;

Query Match 94.6%; Score 229; DB 21; Length 372;
Best Local Similarity 92.7%; Pred. No. 2.8e-20;
Matches 38; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SWPVGPDADICCCNKGDCGTPSHSRPHVMQWSRSYS 41
|||||

DB 308 swfpevgpadicgccknkgdcgtpshrrtqphvmsqwtssas 348

RESULT 4

AA066552
ID AAR66552 standard; Protein; 223 AA.

AC AAR66552;

XX 03-FEB-1995 (first entry)

XX Cynomolgus Monkey zona pellucida ZPC protein.

XX Cynomolgus Monkey; zona pellucida; ZPC; immunoreception.

XX Macaca cynomolgus.

XX MO9411019-A.

XX 26-MAY-1994.

XX 06-NOV-1993; 93WO-US10851.

XX 09-NOV-1992; 92US-0973341.

XX 29-JAN-1993; 93US-0012990.

XX (ZONA-) ZONAGEN INC.

XX Harris JD, Hsu KT, Podolski JS;

XX WPI: 1994-183156/22.

XX N-PSDB; AA079682.

XX Use of zona pellucida proteins and antibodies - for inducing

XX reproducible transient infertility or permanent sterility in

XX female mammals

XX Example 12; Page 132-133; 154pp; English.

XX Cynomolgus monkey cDNA libraries were constructed in lambda gt10

XX using mRNA isolated from a set of ovaries collected from monkeys

XX aged 1.5 and 2 years and a second set from monkeys aged 3, 4 and 14

XX years of age. The libraries were screened with probes encoding

XX porcine ZPA, ZPB and ZPC proteins. Positive clones were analysed

XX further by Southern hybridisation using the porcine probes. Clones

XX encoding cynomolgus monkey ZPA, ZPB and ZPC proteins were

XX identified (see AA079680-079682). AA079682 is the insert from the

XX largest partial ZPC clone which contains just over 50% of the

XX C-terminal portion of the full-length sequence and contains an ORF

XX of 672bp; AAR6552 is the amino acid sequence deduced from the ORF.

XX Sequence 223 AA;

XX

XX

XX

XX

XX

XX

KM Contraceptive: porcine; pig; zona pellucida; mammal; lapine; canine;

KM dog; rabbit; feline; cat; bovine; cow; human; cynomolgus monkey; ZPA;

KM ZPB; ZPC; infertility; veterinary.

OS Macaca fascicularis.

XX US5981228-A.

XX 09-NOV-1999.

XX 07-JUN-1995; 95US-0484596.

XX 09-NOV-1993; 93US-0149223.

XX 09-NOV-1992; 92US-0973341.

XX 29-JAN-1993; 93US-0012990.

XX (ZONA-) ZONAGEN INC.

XX Podolski JS, Hsu KT, Harris JD;

XX WPI: 1999-633318/54.

XX N-PSDB; AA022740.

XX Mammalian zona pellucida proteins used to induce transient or permanent

XX infertility -

XX Example 12; Column 143-146; 84pp; English.

XX This sequence represents the cynomolgus monkey zona pellucida ZPC

XX protein. The invention relates to the isolation of novel nucleotide

XX sequences encoding zona pellucida proteins from mammalian, especially

XX porcine, lapine, canine, feline, bovine, human or cynomolgus monkey

XX sources. The zona pellucida proteins (ZPA, ZPB, and ZPC) are used in

XX the induction of transient or permanent infertility. At present the

XX method is used in veterinary applications to induce transient or

XX permanent infertility in porcine, lapine, canine, feline, bovine,

XX and cynomolgus monkeys.

XX Sequence 223 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```

XX 07-JUN-1995; 95US-0484993.
PF 09-NOV-1993; 93US-0149223.
XX 09-NOV-1992; 92US-0973341.
PR 29-JAN-1993; 93US-0012990.
XX 07-JUN-1995; 95US-0484993.
PA (ZONA-) ZONAGEN INC.
XX Harris JD;
PI WPI: 1999-023447/02.
DR N-PSDB; AAV64820.
XX Isolated zona pellucida DNA from different mammals - used to develop
PT products which can be used for vaccination to induce transient
PT infertility or permanent sterility in female mammals
XX Example 12; Column 145-148; 84pp; English.
XX This sequence represents a Macaca cynomolgus ZPC protein isolated from
CC zona pellucida. This protein can be used in a method for specifically
CC inducing transient infertility or permanent sterility in a host
CC animal by selective vaccination with specific zona pellucida proteins
CC or immunoontraceptively active fragments.
XX Sequence 223 AA:
SQ
Query Match 93.4%; Score 226; DB 20; Length 223;
Best Local Similarity 90.2%; Pred. No. 4e-20;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 SMFPVQGPADICCCCKNGDGPSSRRPPIHVMQNSRSYS 41
107 swfvegpadiqccksgdgtptshrrqphvsgwsrsas 147
RESULT 7
AA82218
ID AAY82218 standard; Protein; 223 AA.
XX AC AAY82218;
XX 12-JUN-2000 (first entry)
XX DE Monkey zona pellucida ZPC SEQ ID NO:49.
XX KW Zona pellucida; ZPA; ZPB; ZPC; infertility; permanent sterility;
XX vaccine; immunoontraceptive; contraceptive; immunostimulatory.
XX OS Primate.
XX US6027727-A.
XX 22-FEB-2000.
XX 09-NOV-1993; 93US-0149223.
XX 09-NOV-1992; 92US-0973341.
XX 29-JAN-1993; 93US-0012990.
XX (ZONA-) ZONAGEN INC.
XX Podolski JS, Hsu KT, Harris JD;
XX WPI: 2000-269144/23.
XX N-PSDB; AA295677.
XX Inducing reproductible transient or permanent infertility in a mammal
PT comprises administration of homologous and/or heterologous mammalian
PT species zona pellucida proteins -

```

```

XX Example 12; Column 147-148; 85pp; English.
PS The present invention describes a method for inducing reproductible
XX transient infertility in a female mammal, including humans, by selective
CC administration of homologous and/or heterologous mammalian species
CC zona pellucida (ZP) proteins or immunoontraceptively active fragments.
CC Also described are: (1) a method for inducing transient infertility in a
CC mammal by administering directly an antibody directed to a ZPB or an
CC immunologically active fragment selected from feline, bovine,
CC cynomolgus monkey or human ZPB polypeptides; (2) an isolated, purified
CC recombinant feline, bovine, cynomolgus monkey or human ZPB polypeptide
CC or immunologically active fragment; and (3) a fusion protein comprising
CC a ZPB polypeptide which is conjugated with a compound selected from
CC keyhole limpet haemocyanin, muramyl dipeptide, histidine-tag, beta-gal,
CC or palmitic acid where the fusion protein remains effective to stimulate
CC production of antibodies that recognise a ZPB polypeptide. The method is
CC useful for providing transient or permanent infertility or sterility in
CC humans and animals. The present sequence represents monkey ZPC, which is
CC used in the exemplification of the present invention.
XX Sequence 223 AA:
SQ
Query Match 93.4%; Score 226; DB 21; Length 223;
Best Local Similarity 90.2%; Pred. No. 4e-20;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 SMFPVQGPADICCCCKNGDGPSSRRPPIHVMQNSRSYS 41
107 swfvegpadiqccksgdgtptshrrqphvsgwsrsas 147
RESULT 8
AAY52692
ID AAY52692 standard; Protein; 223 AA.
XX AC AAY52692;
XX 14-MAR-2000 (first entry)
XX DE Cynomolgus monkey oocyte zona pellucida protein ZPC.
XX KW Zona pellucida; ZPA; ZPB; ZPC; purified; mammalian; glycoprotein;
XX target; immunoontraception; vaccine; antibody; transient; infertility;
XX controllable; predictable; permanent; sterility; side effect.
XX OS Macaca cynomolgus.
XX US6001599-A.
XX 14-DEC-1999.
XX 02-JUN-1995; 95US-0458731.
XX 09-NOV-1993; 93US-0149223.
XX 09-NOV-1992; 92US-0973341.
XX 29-JAN-1993; 93US-0012990.
XX (ZONA-) ZONAGEN INC.
XX Podolski JS, Hsu KT, Harris JD;
XX WPI: 2000-061880/05.
XX Isolated DNA encoding mammalian zona pellucida proteins useful for
PT inducing transient sterility -
XX Example 12; Columns 145-146; 84pp; English.
XX This sequence represents a mammalian zona pellucida protein, ZPC.
CC The zona pellucida is a complex matrix surrounding the mammalian
CC oocyte, formed of glycoproteins secreted by ovarian cells. The

```

invention relates to zona pellucida proteins ZPA, ZPB and ZPC, which are useful as targets for immunocontraceptive vaccines. Polynucleotides encoding these proteins can be used for recombinant protein production, and as probes in hybridisation methods for the isolation of polynucleotides encoding zona pellucida proteins from other mammalian species. Administration of zona pellucida proteins to a host animal results in a specific immunocontraceptive effect. Administration of purified ZPA and/or ZPB induces transient infertility in females. Fertility can be maintained in an on or off condition in a controllable/predictable fashion. Administration of ZPC induces permanent sterility. Use of a purified zona pellucida protein rather than a complex mixture of zona pellucida proteins reduces the potential for unwanted side-effects which may be harmful.

Sequence 223 AA;

Query Match 93.4%; Score 226; DB 21; Length 223;
Best Local Similarity 90.2%; Pred. No. 4e-20; 1; Indels 0; Gaps 0;
Matches 37; Conservative 3; Mismatches 1;

1 SWFPVGGPADICQCCNKDCGTPSHRQPHVMSQMSRSVS 41
||||:|||||||:|||||||:|||||||:|||||||
Db 107 swfpvegpadicqccskgdcgtpshrrqphvsgwsrsas 147

RESULT 9

AA52183
ID AAY52183 standard; Protein: 223 AA.

AC AAY52183;

D7 09-FEB-2000 (first entry)

X2 Monkey zona pellucida protein C (ZPC) amino acid sequence.

X1 Zona pellucida protein: ZPA; ZPB; ZPC; oocyte; ovarian cell; antibody;

KW fusion protein; vaccine; contraceptive; fertilization; fertility;

X1 immunocontraceptive.

O5 Macaca cynomolgus.

PN US5989550-A.

PD 23-NOV-1999.

PF 07-JUN-1995; 95US-0480150.

PR 09-NOV-1993; 93US-0149223.

PR 09-NOV-1992; 92US-0973341.

PR 29-JAN-1993; 93US-0012990.

PA (ZONA-) ZONAGEN INC.

PI Podolski JS, Hsu KT, Harris JD;

DR WPI: 2000-022762/02.

DR N-PSDB; AA237824.

PT Isolated and purified mammalian zona pellucida proteins useful in

XX methods of immunocontraception -

XX Example 12; Column 145-146; 84pp; English.

XX Sequences AAY52181-Y52183 are monkey zona pellucida proteins ZPA, ZPB,

XX and ZPC. The invention relates to isolated and purified zona pellucida

XX proteins from mammals, specifically canine, feline and bovine ZPA or

XX their biologically active fragments also porcine and cynomolgus monkey

XX ZPA+. The zona pellucida is a complex matrix surrounding the mammalian

XX oocyte, formed of glycoproteins secreted by ovarian cells. The

XX previously mentioned ZPA proteins can be used to produce a fusion

XX protein to stimulate production of antibodies in a mammal that recognize

XX a ZPA polypeptide of the mammal. The ZPA polypeptides are useful as

CC vaccines for selectively effecting transient infertility in mammals. ZP

CC has an important role in maintaining the oocyte and in oocyte-sperm

CC interactions and so is a target for a contraceptive agent which

CC interferes with the fertilization process. Providing a specific

CC polypeptide reduces the need for a complex mixture of many ZP proteins

CC which may create unwanted and harmful side effects. The duration of

CC transient fertility is controllable and can be maintained in a

CC controllable and/or predictable fashion.

XX Sequence 223 AA;

Query Match 93.4%; Score 226; DB 21; Length 223;
Best Local Similarity 90.2%; Pred. No. 4e-20; 1; Indels 0; Gaps 0;
Matches 37; Conservative 3; Mismatches 1;

1 SWFPVGGPADICQCCNKDCGTPSHRQPHVMSQMSRSVS 41
||||:|||||||:|||||||:|||||||:|||||||
Db 107 swfpvegpadicqccskgdcgtpshrrqphvsgwsrsas 147

RESULT 10

AA52987
ID AAY52987 standard; Protein: 223 AA.

AC AAY52987;

D7 14-FEB-2000 (first entry)

DE Cynomolgus monkey zona pellucida ZPC protein sequence.

KW Zona pellucida; ZPA; ZPB; ZPC; infertility; sterility; vaccine;

KW immunocontraceptive; contraception; immunisation.

OS Mammalia.

PN US5976545-A.

PD 02-NOV-1999.

PF 07-JUN-1995; 95US-0484158.

PR 09-NOV-1993; 93US-0149223.

PR 09-NOV-1992; 92US-0973341.

PR 29-JAN-1993; 93US-0012990.

PA (ZONA-) ZONAGEN INC.

PI Podolski JS, Hsu KT, Harris JD;

DR WPI: 2000-037080/03.

DR N-PSDB; AA233274.

PT New recombinant mammalian zona pellucida C proteins, for induction of

PT female sterility -

XX Claim 14; Column 145-146; 87pp; English.

XX The present invention describes recombinant zona pellucida (ZP)

XX proteins, specifically ZPC proteins. Also described are fusion proteins

XX of ZPC with any of keyhole limpet haemocyanin, muramyl dipeptide,

XX histidine tag, beta-galactosidase or palmitic acid, capable of

XX stimulating mammalian antibodies that recognise mammalian ZPC.

XX Recombinant ZPC proteins are used to immunise animals to induce transient

XX infertility or permanent female sterility, in humans or other mammals.

XX AA233243 to AA233254, AA233270 to AA233274 and AA233285 encode mammalian

XX ZP proteins given in AAY52971 to AAY52988 from the present invention.

XX AA233255 to AA233269 and AA233275 to AA233284 represent oligonucleotides

XX used in the exemplification of the present invention.

XX Sequence 223 AA;

Query Match 93.4%; Score 226; DB 21; Length 223;
 Best Local Similarity 90.2%; Pred. No. 4e-20;
 Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWPVGADICCCCNKGDGTPSHSRQPHVMSQMSRSVS 41

DB 107 swfvegpadiqccksgdctgtpshsrqphvmsqmsrsas 147

RESULT 11

AA222239

ID AAR22239 standard; Protein; 372 AA.

AC AAR22239;

DT 10-JUL-1992 (first entry)

DE Human ZP3.

XX Zona pellucida; antigenic; epitopes; fertilisation; gametes;
 XX Immunoontraception.

OS Homo sapiens.

EH Key Location/Qualifiers

FT Peptide 341..360
 /note="antigenic epitope"

XX MO9203548-A.

XX 05-MAR-1992.

PF 13-AUG-1991; 91WO-EP01538.

PR 27-AUG-1990; 90EP-0202287.

XX (ALKU) AKZO NV.

XX Van Duin M;

XX WPI: 1992-096892/12.

DR N-PSDB; AAQ22997.

XX Human zona pellucida protein ZP3 and fragments - useful as
 XX contraceptive vaccines or diagnostic medical kits

PS Claim 3; Fig 2; 32pp; English.

XX The protein sequence was deduced from the nucleotide sequence of
 CC zona pellucida ZP3 obtd. by screening a human genomic EMBL3 library
 CC with a labelled ZP3 probe. The human gene was found to comprise
 CC 8 exons spread over 20 kb genomic DNA. cDNA clones were isolated
 CC and expressed in CHO cells. Elucidation of the ZP3 sequence will
 CC allow prodn. of large amts. of the protein by recombinant techniques.
 CC The protein and derived polypeptides may be used to prep. an immuno-
 CC contraction by either active or passive immunisation. The
 CC polypeptide and antibodies (and anti-idiotypic antibodies) directed
 CC to it may be useful in diagnostic kits.

XX Sequence 372 AA;

Query Match 93.0%; Score 225; DB 13; Length 372;

Best Local Similarity 92.5%; Pred. No. 8.7e-20;

Matches 37; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 WFPVGGPADICCCCNKGDGTPSHSRQPHVMSQMSRSVS 41

DB 309 wfpvegpadiqccksgdctgtpshsrqphvmsqmsrsas 348

RESULT 12

AA222239

AA222239

ID AAR53498 standard; Protein; 424 AA.

XX AAR53498;

DT 02-NOV-1994 (first entry)

DE Marmoset ZP3.

XX Zona pellucida; glycoprotein 3; sperm receptor; immunocontraceptive;
 XX diagnosis; tumour; autoimmune disease.

OS Callithrix jacchus.

XX WO9410304-A.

XX 11-MAY-1994.

PF 29-OCT-1993; 93WO-EP03014.

PR 02-NOV-1992; 92EP-0310026.

XX (AKZU) AKZO NOBEL NV.

XX Aitken RJ, Kootan T;

XX WPI: 1994-167461/20.

DR N-PSDB; AAQ63673.

XX New polypeptide(s) with marmoset ZP3 activity - are useful as
 XX immuno-contraceptive vaccines and in diagnosis, also test system
 XX for contraceptive development

XX Claim 1; Page 15; 27pp; English.

XX The marmoset ZP3 gene was isolated by PCR from total marmoset RNA
 CC using PCR primers corresp. to regions of the human ZP3 gene. ZP3
 CC (the sperm receptor) is useful as an immunocontraceptive and as a
 CC diagnostic reagent e.g. to detect sperm function, to monitor the
 CC effects of vaccination, for imaging follicles, to detect ZP3-related
 CC autoimmune disorders or ZP3-expressing tumours or to screen for
 CC autoantibodies against ZP3.

XX Sequence 424 AA;

Query Match 79.3%; Score 192; DB 15; Length 424;

Best Local Similarity 82.1%; Pred. No. 1e-15;

Matches 32; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 SWPVGADICCCCNKGDGTPSHSRQPHVMSQMSRS 39

DB 308 swfvegpadiqccksgdctgtpshsrqphvmsqmsrsas 346

RESULT 13

AA222239

AA222239

AC AAR53199;

DT 31-JAN-1995 (first entry)

DE Canine zona pellucida ZPC protein.

XX Dog; canine; zona pellucida; ZPC; immunocontraception.

XX Canis familiaris.

OS Key Location/Qualifiers

FT Protein 1..426
 /label= canine_ZPC

XX WO9411019-A.

```

PT Mammalian zona pellucida proteins used to induce transient or permanent
PR infertility -
PS Example 3; Column 73-76; 84pp; English.
XX
CC This sequence represents the canine zona pellucida ZPC protein. The
CC invention relates to the isolation of novel nucleotide sequences encoding
CC zona pellucida proteins from mammalian, especially porcine, lapine,
CC canine, feline, bovine, human or cynomolgus monkey sources. The zona
CC pellucida proteins (ZPA, ZPB, and ZPC) are used in the induction of
CC transient or permanent infertility. At present the method is used in
CC veterinary applications to induce transient or permanent infertility
CC in porcine, lapine, canine, feline, bovine, and cynomolgus monkeys.
XX
SQ Sequence      426 AA;

Query Match          62.4%; Score 151; DB 20; Length 426;
Best Local Similarity 60.0%; Pred. No. 1e-10;
Matches    24; Conservative   5; Mismatches    11; Indels     0; Gaps     0;

Oy       2 WEPVGGPADICGCCNKGDCTPSSHRQDPHWSGSRSSVS 41
Db        |||::|||:||||| || | ||| : : |||
           307 wyvegsadictccnksgcglpgrrishlergwrksvs 346

RESULT 15
ID AAWB1809 standard; Protein: 426 AA.
AC AAWB1809;
DT 29-JAN-1999 (first entry)
DE Canine ZPC protein.
KW ZPC; zona pellucida; infertility; sterility; immunoncontraceptive;
  vaccine; canine.
OS Canis sp.
PN US5637497-A.
PD 17-NOV-1998.
PE 07-JUN-1995; 95US-0484993.
PF 09-NOV-1993; 93US-0149223.
PR 09-NOV-1992; 92US-0973341.
PP 29-JAN-1993; 93US-0012990.
PX 07-JUN-1995; 95US-0484993.
PY (ZONA-) ZONAGEN INC.
PI Harris JD;
DR MPI; 1999-023447/02.
DN N-PSTDB; AAV64791.
EX Isolated zona pellucida DNA from different mammals - used to develop
  products which can be used for vaccination to induce transient
  infertility or permanent sterility in female mammals.
PS Claim 5; Column 78-80; 84pp; English.
CC This sequence represents a canine ZPC protein isolated from zona
CC pellucida. This protein can be used in a method for specifically
CC inducing transient infertility or permanent sterility in a host
CC animal by selective vaccination with specific zona pellucida proteins
CC or immunoncontraceptively active fragments.
XX
SQ Sequence      426 AA;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 12, 2001, 08:24:10 ; Search time 24.08 Seconds
(without alignments)
34.299 Million cell updates/sec

Title: US-09-252-828A-2
Perfect score: 242
Sequence: 1 SMFVPGRADICQCCNKGDG.....TPSHRRQPHVMSQMSRVS 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues
Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents,AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	235	97.1	424	1	US-08-453-472-7 Sequence 7, Appl
2	235	97.1	424	1	US-08-038-948-4 Sequence 4, Appl
3	235	97.1	424	1	US-08-453-952-7 Sequence 7, Appl
4	235	97.1	424	2	US-08-862-903-7 Sequence 7, Appl
5	235	97.1	424	2	US-08-484-158B-61 Sequence 61, Appl
6	226	93.4	223	2	US-08-484-993B-49 Sequence 49, Appl
7	226	93.4	223	2	US-08-484-158B-49 Sequence 49, Appl
8	226	93.4	223	2	US-08-484-596A-49 Sequence 49, Appl
9	226	93.4	223	3	US-08-480-150A-49 Sequence 49, Appl
10	226	93.4	223	3	US-08-458-731-49 Sequence 49, Appl
11	226	93.4	223	3	US-08-149-223A-49 Sequence 49, Appl
12	226	93.4	426	2	US-08-484-993B-12 Sequence 12, Appl
13	226	93.4	426	2	US-08-484-158B-12 Sequence 12, Appl
14	226	93.4	426	2	US-08-484-596A-12 Sequence 12, Appl
15	226	93.4	426	2	US-08-480-150A-12 Sequence 12, Appl
16	226	93.4	426	3	US-08-458-731-12 Sequence 12, Appl
17	226	93.4	426	3	US-08-149-223A-12 Sequence 12, Appl
18	226	93.4	426	2	US-08-484-993B-18 Sequence 18, Appl
19	226	93.4	426	2	US-08-484-158B-18 Sequence 18, Appl
20	226	93.4	426	2	US-08-484-596A-18 Sequence 18, Appl
21	226	93.4	426	2	US-08-480-150A-18 Sequence 18, Appl
22	226	93.4	426	3	US-08-458-731-18 Sequence 18, Appl
23	226	93.4	426	3	US-08-149-223A-18 Sequence 18, Appl
24	226	93.4	426	1	US-08-453-472-8 Sequence 8, Appl
25	226	93.4	426	1	US-08-038-948-2 Sequence 5, Appl
26	226	93.4	426	1	US-08-038-948-5 Sequence 5, Appl
27	226	93.4	426	1	US-08-453-952-8 Sequence 8, Appl

28	126	52.1	424	2	US-08-862-903-8	Sequence 8, Appl
29	119	49.2	421	2	US-08-484-993B-6	Sequence 6, Appl
30	119	49.2	421	2	US-08-484-158B-6	Sequence 6, Appl
31	119	49.2	421	2	US-08-484-596A-6	Sequence 6, Appl
32	119	49.2	421	2	US-08-480-150A-6	Sequence 6, Appl
33	119	49.2	421	3	US-08-458-731-6	Sequence 6, Appl
34	119	49.2	421	3	US-08-149-223A-6	Sequence 6, Appl
35	109	45.0	421	2	US-08-484-993B-24	Sequence 24, Appl
36	109	45.0	421	2	US-08-484-158B-24	Sequence 24, Appl
37	109	45.0	421	2	US-08-484-596A-24	Sequence 24, Appl
38	109	45.0	421	2	US-08-480-150A-24	Sequence 24, Appl
39	109	45.0	421	3	US-08-458-731-24	Sequence 24, Appl
40	109	45.0	421	3	US-08-149-223A-24	Sequence 24, Appl
41	100	41.3	415	2	US-08-484-993B-8	Sequence 8, Appl
42	100	41.3	415	2	US-08-484-158B-8	Sequence 8, Appl
43	100	41.3	415	2	US-08-484-596A-8	Sequence 8, Appl
44	100	41.3	415	2	US-08-480-150A-8	Sequence 8, Appl
45	100	41.3	415	3	US-08-458-731-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-453-472-7
Sequence 7, Application US/08453472
Patent No. 5626846
GENERAL INFORMATION:
APPLICANT: DEAN, JURRIEN
TITLE OF INVENTION: CONTRACEPTIVE VACCINE
TITLE OF INVENTION: BASED ON ALLOIMUNIZATION WITH ZONA PELLUCIDA
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,472
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/038,948
FILING DATE: 26-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,462
FILING DATE: 20-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/364,379
FILING DATE: 12-JUN-1989
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4032 US3
TELEPHONE: (212) 751-6849
TELEFAX: (212) 751-6849
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 424
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein

```

ORIGINAL SOURCE:
ORGANISM: human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY: ZP3
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: human ZP3 protein
US-08-453-472-7

Query Match 97.1%; Score 235; DB 1; Length 424;
Best Local Similarity 95.1%; Pred. No. 6.4e-21;
Matches 39; Conservatively 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SMPVPGPADICCCCKNGDCGPHSRSRPHYMSQMSRSYS 41
Db 308 SMPVEGPADICCCCKNGDCGPHSRSRPHYMSQMSRSAS 348
|||||

RESULT 2
US-08-038-948-4
Sequence 4, Application US/08038948
Patent No. 5641487
GENERAL INFORMATION:
APPLICANT: DEAN, JURRIEN
TITLE OF INVENTION: CONTRACEPTIVE VACCINE BASED ON
TITLE OF INVENTION: ALLOIMMUNIZATION WITH ZONA PELLUCIDA POLYPEPTIDES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/038,948
FILING DATE: 26-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/930,462
FILING DATE: 20-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/364,379
FILING DATE: 12-JUN-1989
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: 99152/E-266-88/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

```

```

US-08-038-948-4
Query Match          97.1%; Score 235; DB 1; Length 424;
Best Local Similarity 95.1%; Pred. No. 6.4e-21;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1  SWFPGVGPADICCCCNKGDCGTPSHSRROPHYMSQMSRSVS 41
        |||F|||||
DB      308  SWFPGVGPADICCCCNKGDCGTPSHSRROPHYMSQMSRSAS 348

RESULT      3
US-08-453-952-7
; Sequence 7, Application US/08453952
; Patent No. 5672488
; GENERAL INFORMATION:
; APPLICANT: DEAN, JURRIEN
; TITLE OF INVENTION: CONTRACEPTIVE VACCINE
; TITLE OF INVENTION: BASED ON ALLOIMMUNIZATION WITH ZONA PELLUCIDA
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,952
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/038,948
; FILING DATE: 26-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,462
; FILING DATE: 20-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/364,379
; FILING DATE: 12-JUN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4032 US4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPOLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; FEATURE:

```

```

:      TOPOLOGY: linear
:
:      MOLECULE TYPE: protein
:
:      DESCRIPTION:
:      /desc =
:      DESCRIPTION: "deduced amino acid sequence of human ZPC
:

```

US-08-484-158B-61

Query Match 97.18; Score 235; DB 2; Length 424;
Best Local Similarity 95.18; Pred. No. 6.4e-21;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWFVVGPDADICCCNKGDCGTPSHSRPHVMSQMSRSVS 41
DB 308 SWFVVGPDADICCCNKGDCGTPSHSRPHVMSQMSRSVS 348

RESULT 6

US-08-484-993B-49
Sequence 49, Application US/08484993B
Patent No. 5837497
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,993B
FILING DATE: 09-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-993B-49

Query Match 93.48; Score 226; DB 2; Length 223;
Best Local Similarity 90.28; Pred. No. 4.1e-20;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWFVVGPDADICCCNKGDCGTPSHSRPHVMSQMSRSVS 41
DB 107 SWFVVGPDADICCCNKGDCGTPSHSRPHVMSQMSRSVS 147

RESULT 7

US-08-484-158B-49
Sequence 49, Application US/08484158B

Patent No. 5976545

GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Pharmaceutical Compositions for
Immunocontraception
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,158B
FILING DATE: 07-JUNE-95
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 09-NOV-93
APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-92
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 32794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-158B-49

Query Match 93.48; Score 226; DB 2; Length 223;
Best Local Similarity 90.28; Pred. No. 4.1e-20;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWFVVGPDADICCCNKGDCGTPSHSRPHVMSQMSRSVS 41
DB 107 SWFVVGPDADICCCNKGDCGTPSHSRPHVMSQMSRSVS 147

RESULT 8

US-08-484-596A-49
Sequence 49, Application US/08484596A
Patent No. 5981228

GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,596A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,223
FILING DATE: 11-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-596A-49

Query Match 93.4%; Score 226; DB 2; Length 223;
Best Local Similarity 90.2%; Pred. No. 4.1e-20;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWFVPGPADICCCCKNGDCTPSHSRQPHVMSQMSRSVS 41
|||||
Db 107 SWFVPGPADICCCCKNGDCTPSHSRQPHVMSQMSRSVS 147

RESULT 9
US-08-480-150A-49
Sequence 49, Application US/08480150A
Patent No. 5989550
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,150A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/149,223

FILING DATE: 09-NOV-1993
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-150A-49

Query Match 93.4%; Score 226; DB 2; Length 223;
Best Local Similarity 90.2%; Pred. No. 4.1e-20;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWFVPGPADICCCCKNGDCTPSHSRQPHVMSQMSRSVS 41
|||||
Db 107 SWFVPGPADICCCCKNGDCTPSHSRQPHVMSQMSRSVS 147

RESULT 10
US-08-458-731-49
Sequence 49, Application US/08458731
Patent No. 6001599
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,731
FILING DATE: 09-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-458-731-49

Query Match 93.4%; Score 226; DB 3; Length 223;
Best Local Similarity 90.2%; Pred. No. 4.1e-20;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWPVGSPADICCCCKGDCGTPSHSRQPHVMSQMSRSYS 41
Db 107 SWPVGSPADICCCCKGDCGTPSHSRQPHVMSQMSRSYS 147

RESULT 11
US-08-149-223A-49
Sequence 49, Application US/08149223A
Patent No. 6027727
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149, 223A
FILING DATE: 09-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELETYPE: 25-3856
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-149-223A-49

Query Match 93.4%; Score 226; DB 3; Length 223;
Best Local Similarity 90.2%; Pred. No. 4.1e-20;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
OY 1 SWPVGSPADICCCCKGDCGTPSHSRQPHVMSQMSRSYS 41
Db 107 SWPVGSPADICCCCKGDCGTPSHSRQPHVMSQMSRSYS 147

RESULT 12
US-08-484-993B-12
Sequence 12, Application US/08484993B
Patent No. 5837497
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Materials and Methods for Immunocontraception
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,993B
FILING DATE: 09-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/012,990
FILING DATE: 29-JAN-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/973,341
FILING DATE: 09-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 31745
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6653
TELEFAX: 312/474-0448
TELETYPE: 25-3856
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-993B-12

Query Match 62.4%; Score 151; DB 2; Length 426;
Best Local Similarity 60.0%; Pred. No. 7.5e-11;
Matches 24; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

OY 2 WFPVGSADICCCCKGDCGTPSHSRQPHVMSQMSRSYS 41
Db 307 WFPVGSADICCCCKGDCGTPSHSRQPHVMSQMSRSYS 346

RESULT 13
US-08-484-158B-12
Sequence 12, Application US/08484158B
Patent No. 5976545
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Jeffrey D.
APPLICANT: Hsu, Kuang T.
APPLICANT: Podolski, Joseph S.
TITLE OF INVENTION: Pharmaceutical Compositions for
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

04 protein - protein search, using sw model

Run on: July 12, 2001, 08:24:10 ; Search time 41.53 Seconds
(without alignments)
75.202 Million cell updates/sec

Title: US-09-252-828a-2

Sequence: 1 SWFPGGADICCCCKNGDC.....TSHSRQPHVMSQMSRSVS 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 segs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	235	97.1	424	1	A36000	sperm-binding glyco
2	217	89.7	210	2	A56844	POM-zp3 protein -
3	151	62.4	426	2	S70396	zona pellucida gly
4	143	59.1	432	1	A60503	sperm-binding glyco
5	136	56.2	424	2	S70399	zona pellucida gly
6	126	52.1	424	1	A30334	sperm-binding glyco
7	119	49.2	421	1	S70433	zona pellucida gly
8	109	45.0	421	1	S70402	zona pellucida gly
9	100	41.3	415	2	S70401	zona pellucida gly
10	98	40.5	44	2	B4365	sperm receptor lig
11	60.5	25.0	99	2	B82720	hypothetical prote
12	59	24.4	58	2	S35373	zona pellucida gly
13	58.5	24.4	999	1	IJH063	desmoglein 3 precu
14	58	24.2	357	2	S09269	ig alpha chain C r
15	56.5	23.3	66	2	S78705	protein YBR056w-a
16	56	23.1	819	2	G81698	leucyl-tRNA synthet
17	55.5	22.9	647	2	JE0337	Frizzled-1 protein
18	55	22.7	436	2	JN0591	serotonin receptor
19	55	22.7	641	2	A45054	probable intercell
20	55	22.7	858	2	T08881	prominin - mouse
21	55	22.5	865	2	T09050	AC133 antigen - hu
22	54.5	22.5	424	2	S52847	egg membrane prote
23	54.5	22.5	735	2	I48101	ADAM 6 protein pre
24	54	22.3	202	2	A33176	P21 protein - soyb
25	54	22.3	477	2	A47236	zinc-finger protei
26	54	22.3	494	2	A42170	zinc finger protei
27	54	22.3	497	2	JC5076	myc-associated zin
28	53.5	22.1	351	2	C72508	hypothetical prote
29	53.5	22.1	379	2	H96696	protein FIN21.16 l

30	53.5	22.1	738	2	T00343	hypothetical prote
31	53	21.9	197	2	I46413	keratin KAP5.5 - s
32	53	21.9	563	2	T47520	hypothetical prote
33	52.5	21.7	447	2	S39505	translatlon elonga
34	52.5	21.7	447	2	S17434	translatlon elonga
35	52.5	21.7	447	2	JS0719	translatlon elonga
36	52.5	21.7	447	2	JC1454	translatlon elonga
37	52.5	21.7	447	2	S66339	translatlon elonga
38	52.5	21.7	448	2	S10507	translatlon elonga
39	52.5	21.7	449	2	S08534	translatlon elonga
40	52.5	21.7	449	2	S06724	translatlon elonga
41	52.5	21.7	449	2	S21989	translatlon elonga
42	52.5	21.7	967	2	F86214	protein T6D22.2 [i
43	52	21.5	127	2	T08097	osmotin - sweet or
44	51.5	21.3	339	2	F86614	phenylalanyl tRNA
45	51.5	21.3	339	2	B72011	phenylalanine--CRN

ALIGNMENTS

RESULT 1
A36000
Sperm-binding glycoprotein ZP3 precursor - human
N:Alternate names: sperm receptor ZP3; zona pellucida glycoprotein ZP3
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text-change 10-Sep-1999
C:Accession: A36000; A44365
R:Chamberlin, M.E.; Dean, J.
Proc. Natl. Acad. Sci. U.S.A. 87, 6014-6018, 1990
A:Title: Human homolog of the mouse sperm receptor.
A:Reference number: A36000; MUID:90349545
A:Accession: A36000
A:Molecule type: mRNA; DNA
A:Residues: 1-424 <CHA>
A:Cross-references: GB:M60504; GB:M35109; NID:9340491; PIDN:AAA61336.1; PID:9340492
R:van Duin, M.; Polman, J.E.; Verkoelen, C.C.; Bunschoten, H.; Meyerink, J.H.; Olijve
Genomics 14, 1064-1070, 1992
A:Title: Cloning and characterization of the human sperm receptor ligand ZP3: evidenc
A:Reference number: A44365; MUID:93122771
A:Accession: A44365
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 329-370, 'S', 372-424 <VAN>
A:Experimental source: ovary
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBI:122391, NCBI:122392)
C:Comment: This sulfated glycoprotein in the zona pellucida of the oocyte is a recept
C:Genetic: GDB:2P3A
A:Gene: GDB:128007; OMIM:182889
A:Cross-references: GDB:128007; OMIM:182889
A:Map position: 7pter-7qter
C:Keywords: sperm-binding glycoprotein ZP3; ZP domain homology
C:Keywords: glycoprotein; oocyte; receptor; sulfolipin; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-424/Product: sperm-binding glycoprotein ZP3 #status predicted <MAT>
F:45-301/Domain: ZP domain homology <ZPH>

Query Match 97.1% Score 235; DB 1; Length 424;
Best Local Similarity 95.1% Pred. No. 1.2e-20;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWFPGGADICCCCKNGDCGTPSHSRQPHVMSQMSRSVS 41
|||||
Db 308 SWFPGGADICCCCKNGDCGTPSHSRQPHVMSQMSRSVS 348

RESULT 2
A56844
POM-zp3 protein - human
C:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence-revision 19-Oct-1995 #text-change 05-Nov-1999

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-152, 'E', 154-252, 'E', 254-424 <ROS>

A:Experimental source: ovary

A:Note: sequence extracted from NCBI backbone (NCBIP:119137)

C:Comment: This sulfated glycoprotein in the zona pellucida of the oocyte is a receptor

C:Genetics:

A:Introns: 103/3; 143/2; 178/1; 239/2; 278/3; 309/2; 355/1

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

C:Keywords: glycoprotein; oocyte; sulfoprotein; transmembrane protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-424/Product: sperm-binding glycoprotein ZP3 #status predicted <MAT>

F:45-302/Domain: ZP domain homology <ZPH>

F:146,304,327,330/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:273/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 52.1%; Score 126; DB 1; Length 424;

Best Local Similarity 53.7%; Pred. No. 1.1e-07;

Matches 22; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 SWPVGADICCCCKNGDCGTPSHSRPHVMSQMSRSVS 41

DB 309 SWPVGADICCCCKNGDCGTPSHSRPHVMSQMSRSVS 349

RESULT 7

zona pellucida glycoprotein C - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

DNA Seq. 4, 361-393, 1994

A:Title: Cloning and characterization of zona pellucida genes and cDNAs from a variety

A:Reference number: S70396; MUID:95143578

A:Accession: S70433

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-421 <HAR>

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

F:44-300/Domain: ZP domain homology <ZPH>

Query Match 49.2%; Score 119; DB 1; Length 421;

Best Local Similarity 70.4%; Pred. No. 7.4e-07;

Matches 19; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 WFPVGPADICCCCKNGDCGTPSHSR 28

DB 308 WSPVEGPADICCCCKNGDCGTPSHSR 334

RESULT 8

zona pellucida glycoprotein C - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S70402

DNA Seq. 4, 361-393, 1994

A:Title: Cloning and characterization of zona pellucida genes and cDNAs from a variety

A:Reference number: S70396; MUID:95143578

A:Accession: S70402

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-421 <HAR>

A:Cross-references: EMBL:U05775; NID:g458266; PIDN:AAA74385.1; PID:g458267

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

F:44-300/Domain: ZP domain homology <ZPH>

Query Match 45.0%; Score 109; DB 1; Length 421;

Best Local Similarity 56.7%; Pred. No. 1.1e-05;

Matches 17; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 2 WFPVGPADICCCCKNGDCGTPSHSRPH 31

DB 308 WSPVEGPTDICCCCKNGDCGTPSHSRPH 337

RESULT 9

zona pellucida glycoprotein C - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 20-Aug-1999

C:Accession: S70401

DNA Seq. 4, 361-393, 1994

A:Title: Cloning and characterization of zona pellucida genes and cDNAs from a variety

A:Reference number: S70396; MUID:95143578

A:Accession: S70401

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-415 <HAR>

A:Cross-references: EMBL:U05782; NID:g458280; PIDN:AAA74392.1; PID:g458281

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

F:41-295/Domain: ZP domain homology <ZPH>

Query Match 41.3%; Score 100; DB 2; Length 415;

Best Local Similarity 42.6%; Pred. No. 0.00013;

Matches 20; Conservative 5; Mismatches 10; Indels 12; Gaps 2;

OY 1 SWPVGADICCCCKNGDC-----GTP-----SHSRPHVMSQ 35

DB 302 SWPVGADICCCCKNGDCDLIAGSPNQNHRARSLSRRHVTVE 348

RESULT 10

B44365

sperm receptor ligand homolog ZP3-372 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 12-May-1995

C:Accession: B44365

R:van Duijn, M.; Polman, J.F.; Verkoelen, C.C.; Bunschoten, H.; Meyerink, J.H.; Olthoff

Genomics 14, 1064-1070, 1992

A:Title: Cloning and characterization of the human sperm receptor ligand ZP3: evidence

A:Reference number: A44365; MUID:93122771

A:Accession: B44365

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-44 <VAN>

A:Experimental source: ovary

A:Note: sequence extracted from NCBI backbone (NCBIN:123950, NCBIP:122393)

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

Query Match 40.5%; Score 98; DB 2; Length 44;

Best Local Similarity 90.0%; Pred. No. 4e-05;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 22 TPSSRRPHVMSQMSRSVS 41

DB 1 TPSSRRPHVMSQMSRSVS 20

RESULT 11

B82720

hypothetical protein xrl139 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: B82720

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: AB2515; MUID:20365717
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: B82720
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-99 <SIM>
A:Cross-references: GB:AE003949; GB:AE003849; NID:99106088; PIDN:AAF83949.L; GSPDB:GM000000
R:Experimental source: strain 985C
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Bittone, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, I.E.A.; Carraro, D.M.; Carter, H.; de Nêto, E.; Docena, C.; El-Dorri, H.; Fachinani, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohlich, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitaajima, J.P.; Krieger, J.E.; Krumme, E.E.; Laigle, J.D.; Machado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, A.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Meick, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki, M.; Tsubakio, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zucchi, A.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
C:Gene: Xf1139

A: Reference number: S45906
 A: Accession: S78705
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-66 <AlJ>
 A: Cross-references: EMBL:Z35925; MIPS:YBR056w-a
 A: Experimental source: strain S288C
 C: Genetics:
 A: Map position: 2R

Query Match 24.0%; Score 58; DB 2; Length 66;
 Best Local Similarity 57.9%; Pred. No. 3.1;
 Matches 11; Conservative 0; Mismatches 4; Indels 4; Gaps 1;
 Q7 6 QGPA---DQCQCCKRGDC 20
 ||| | | | | | | |
 D3 24 QGPPPPRNDCCCCCNCGDC 42

Search completed: July 12, 2001, 08:26:25
 Job time: 135 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

CM protein - protein search, using sw model

Run on: July 12, 2001, 08:24:10 ; Search time 10.02 Seconds
(without alignments)
140.167 Million cell updates/sec

Title: US-09-252-828A-2

Sequence: 1 SMFVQGRADICCCCKGDC.....TSHSRQPRHVSQMSRSVS 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	235	97.1	424	1	P21754 homo sapien
2	229	94.6	472	1	P06633 homo sapien
3	226	93.4	424	1	P53785 macaca radi
4	192	79.3	424	1	P53786 callithrix
5	151	62.4	426	1	P48831 canis fami
6	143	59.1	422	1	P23491 mesocricetu
7	136	56.2	424	1	P48832 felis silve
8	126	52.1	424	1	P23_MOUSE
9	119	49.2	421	1	P43098 sus scrofa
10	109	45.0	421	1	P48830 bos taurus
11	100	41.3	415	1	P48833 oryctolagus
12	99	24.4	424	1	P33926 homo sapien
13	55.5	22.9	436	1	P3388 rattus norv
14	55	22.7	865	1	P43490 homo sapien
15	54	22.7	867	1	P25090 mus musculu
16	54	22.3	202	1	P21_SORBN
17	54	22.3	477	1	P53270 homo sapien
18	54	22.3	477	1	P56671 mus musculu
19	53.5	22.1	338	1	O76075 homo sapien
20	52.5	21.7	447	1	EF11_HORVU
21	52.5	21.7	447	1	EF12_DAUCA
22	52.5	21.7	447	1	EF12_HORVU
23	52.5	21.7	447	1	EF1A_MAIZE
24	52.5	21.7	447	1	EF1A_ORYSA
25	52.5	21.7	447	1	EF1A_PEA
26	52.5	21.7	447	1	EF1A_SOYBN
27	52.5	21.7	447	1	EF1A_TOBAC
28	52.5	21.7	447	1	EF1A_VICFA
29	52.5	21.7	447	1	EF1A_WHEAT
30	52.5	21.7	448	1	EF1A_LYCES
31	52.5	21.7	449	1	EF11_DAUCA
32	52.5	21.7	449	1	EF1A_ARATH
33	52.5	21.7	449	1	EF1A_MANES

34	51.5	21.3	339	1	SYEA_CHLUP
35	51	21.1	65	1	TXW6_NAJNA
36	51	21.1	281	1	TRY2_DROBR
37	51	21.1	1132	1	DNEI_HSV6U
38	51	21.1	1132	1	DNEI_HSV6Z
39	50.5	20.9	74	1	NXLD_ACAAN
40	50.5	20.9	139	1	IGE_MYXGL
41	50.5	20.9	622	1	ACHA_CHICK
42	50	20.7	174	1	SODC_CAUCR
43	50	20.7	182	1	KRUC_SHEEP
44	50	20.7	269	1	ILIB_MOUSE
45	50	20.7	643	1	RHOP_MOUSE

ALIGNMENTS

RESULT 1
ID ZP3A_HUMAN STANDARD; PRT: 424 AA.
AC P21754;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A PRECURSOR (ZONA PELLUCIDA
DE GLYCOPROTEIN ZP3A) (ZONA PELLUCIDA PROTEIN C) (SPERM RECEPTOR) (ZP3).
GN ZP3A OR ZP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90349545; PubMed=2385582;
RA Chamberlin M.E.; Dean J.;
RT "Human homolog of the mouse sperm receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 87:6014-6018(1990).
RN [2]
RP SEQUENCE OF 329-424 FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=93122771; PubMed=1478648;
RA van Duin M., Polman J.E., Verkoelen C.C., Bunschoten H.,
RT Meyerink J.H., Olijve W., Aitken R.J.;
RL "Cloning and characterization of the human sperm receptor ligand ZP3:
evidence for a second polymorphic allele with a different frequency
in the Caucasian and Japanese populations."
RL Genomics 14:1064-1070(1992).
CC -!- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
CC Sperm-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
CC SPECIES-SPECIFICITY OF THE INSEMINATION.
CC -!- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
CC MATRIX.
CC -!- PTM: SUFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.
CC -!- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M60504; AAA61336.1; -.
CC PIR: A36000; A36000.
CC MIM: 182889; -.
CC InterPro: IPR001507; -.
CC Pfam: PF00100; zona_pellucida; 1.
CC PRINTS: PR00023; ZPELUCIDA.
CC PROSITE: PS00682; ZP_DOMAIN; 1.
KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;

KW Extracellular matrix; Multigene family.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A.
 FT DOMAIN 23 387 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 388 408 POTENTIAL.
 FT DOMAIN 409 424 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 424 452 ZP.
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 424 AA; 47028 MW; 9451/B7B6E9F06 CRC64;

Query Match 97.1%; Score 235; DB 1; Length 424;
 Best Local Similarity 95.1%; Pred. No. 2e-21;
 Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWFVPGPADICCCCKNGDCGTPSHSRORPHVMSQWSRSVS 41
 DB 308 SWFVPGPADICCCCKNGDCGTPSHSRORPHVMSQWSRSVS 348

RESULT 2
 ZP3B_HUMAN STANDARD; PRT; 372 AA.
 AC 006633;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3B PRECURSOR (ZONA PELLUCIDA
 DE GLYCOPROTEIN ZP3B) (SPERM RECEPTOR) (ZP3).
 GN ZP3B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCBL_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=9312771; PubMed=1478648;
 RA van Duin M., Polman J.E., Verkoelen C.C., Bunschoten H.,
 RA Meyerink J.H., Olijve W., Aitken R.J.;
 RT "Cloning and characterization of the human sperm receptor ligand ZP3:
 RT evidence for a second polymorphic allele with a different frequency
 RT in the Caucasian and Japanese populations.";
 RL Genomics 14:1064-1070(1992).
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
 CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
 CC SPECIES-SPECIFICITY OF THE INSEMINATION.
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
 CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: X56777; CAA40095.1; -;
 CC EMBL: A18567; CAA01398.1; -;
 DR InterPro: IPR001507; -;
 DR Pfam: PF00100; zona_pellucida; 1.
 DR PRINTS: PR00023; ZPELUCIDA.
 DR PROSITE: PS00682; ZP_DOMAIN; 1.
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor;
 KW Extracellular matrix; Multigene family.

FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 372 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3B.
 FT DOMAIN 23 307 ZP.
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 372 AA; 41424 MW; AB78DB5311BE7B4A CRC64;

Query Match 94.6%; Score 229; DB 1; Length 372;
 Best Local Similarity 92.7%; Pred. No. 9.6e-21;
 Matches 38; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SWFVPGPADICCCCKNGDCGTPSHSRORPHVMSQWSRSVS 41
 DB 308 SWFVPGPADICCCCKNGDCGTPSHSRORPHVMSQWSRSVS 348

RESULT 3
 ZP3A_MACRA STANDARD; PRT; 424 AA.
 AC P53785;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A PRECURSOR (ZONA PELLUCIDA
 DE GLYCOPROTEIN ZP3A) (ZONA PELLUCIDA PROTEIN C) (SPERM RECEPTOR) (ZP3).
 GN ZP3A OR ZP3.
 OS Macaca radiata (Bonnet monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 NC NCBL_TaxID=9548;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=96249321; PubMed=8848588;
 RA Koliuri S.K., Kaul R., Banerjee K., Gupta S.K.;
 RT "Nucleotide sequence of cDNA encoding bonnet monkey (Macaca radiata)
 RT Reprod. Fertili. Dev. 7:1209-1212(1995).
 RL Reprod. Fertili. Dev. 7:1209-1212(1995).
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
 CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
 CC SPECIES-SPECIFICITY OF THE INSEMINATION.
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
 CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
 CC MATRIX.
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: X82639; CAA57961.1; -;
 DR InterPro: IPR001507; -;
 DR Pfam: PF00100; zona_pellucida; 1.
 DR PRINTS: PR00023; ZPELUCIDA.
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
 KW Extracellular matrix; Multigene family.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A.
 FT DOMAIN 23 387 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 388 408 POTENTIAL.
 FT DOMAIN 409 424 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 424 452 ZP.

FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 424 AA; 47040 MW; 3B41C4CFA3792331 CRC64;

Query Match 93.4%; Score 226; DB 1; Length 424;
 Best Local Similarity 90.2%; Pred. No. 2,5e-20;
 Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SWFVPGPADICCCCKGDCGTPSHRQPHVMSQMSRSVS 41
 DB 308 SWFVPGPADICCCCKGDCGTPSHRQPHVMSQMSRSVS 348

RESULT 4
 ID ZP3A_CALSO STANDARD; PRT; 424 AA.
 AC P53786;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A PRECURSOR (ZONA PELLUCIDA GLYCOPROTEIN ZP3A) (ZONA PELLUCIDA PROTEIN C) (SPERM RECEPTOR) (ZP3).
 GN ZP3A OR ZP3.
 OS Callithrix sp. (Marmoset).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
 OC NCBI_TaxID=9485;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=94363314; PubMed=8081814;
 RA Thilak-Kootheran P., Van Duin M., Altken R.J.;
 RT "Cloning, sequencing and oocyte-specific expression of the marmoset sperm receptor protein, ZP3.";
 RL Zygote 1:93-101(1993).
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE SPECIES-SPECIFICITY OF THE INSEMINATION.
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR MATRIX.
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: S71825; AAB31866.1; -;
 DR InterPro: IPR001507; -;
 DR Pfam: PF00100; zona_pellucida; 1.
 DR PRINTS: PR00023; ZPELUCIDA.
 DR PROSITE: PS00682; ZP_DOMAIN; 1.
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane; Extracellular matrix; Multigene family;
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A.
 FT DOMAIN 23 387 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 388 408 POTENTIAL.
 FT DOMAIN 409 424 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 45 307 ZP.
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 424 AA; 46809 MW; 1DACBD03026C2739 CRC64;

Query Match 79.3%; Score 192; DB 1; Length 424;
 Best Local Similarity 82.1%; Pred. No. 3.1e-16;
 Matches 32; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 SWFVPGPADICCCCKGDCGTPSHRQPHVMSQMSRS 39
 DB 308 SWFVPGPADICCCCKGDCGTPSHRQPHVMSQMSRS 346

RESULT 5
 ID ZP3_CANFA STANDARD; PRT; 426 AA.
 AC P48831;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
 GN ZP3 OR ZPC.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OC NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=95143578; PubMed=7841460;
 RA Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yvewicz E.C., Sacco A.G.;
 RT "Cloning and characterization of zona pellucida genes and cDNAs from a variety of mammalian species: the ZPA, ZPB and ZPC gene families.";
 RL DNA Seq. 4:361-393(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RA Okazaki Y., Sugimoto M.;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE SPECIES-SPECIFICITY OF THE INSEMINATION (BY SIMILARITY).
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR MATRIX.
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U05780; AAA74387.1; -;
 DR EMBL: DA5070; BAA08098.1; -;
 DR InterPro: IPR001507; -;
 DR Pfam: PF00100; zona_pellucida; 1.
 DR PRINTS: PR00023; ZPELUCIDA.
 DR PROSITE: PS00682; ZP_DOMAIN; 1.
 KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane; Extracellular matrix; Multigene family;
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 426 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
 FT DOMAIN 23 385 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 386 406 POTENTIAL.
 FT DOMAIN 407 426 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 43 305 ZP.

	QY	DB	mismatches	Indels	Gaps
	2 WFFVVGSPADICCCCKNGGCGTPSHSRORPQWMSRVS	307 WYPVESGADICRCCKNGSGCLPGSRRLSLHLEKRWKRVS	3	11	0
	.: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .	.: .: .: .: .: .: .: .: .: .: .: .: .: .: .: .			
	41	41			

01-NOV-1991 (Rel. 20, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
Mesocricetus auratus (Golden hamster).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus.
NCBI_TaxId=10036;
(1)
SEQUENCE FROM N.A.
TISSUE-Ovary;
MEDLINE=91078540; PubMed=2257975;
Kiloch R.A., Ruiz-Seller B., Massaman P.M.;
Genomic organization and polypeptide primary structure of zona Dev. Biol. 142:414-421(1990).
-1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE SPECIES-SPECIFICITY OF THE INSEMINATION.
-1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1. MATRIX.
-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR TISSUE SPECIFICITY: OOCYTES.
-1- DEVELOPMENTAL STAGE: GROWING OOCYTES.
-1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.
-1- SIMILARITY: CONTAINS 1 ZP DOMAIN.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as their content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: M63629, AAA37079.1; -.
InterPro: IPR001507; -.
Pfam: PF00100; zona_pellucida_1.
PRINTS: PR00023; ZPELLUCIDA.
PROSITE: PS00682; ZP_DOMAIN; 1.
Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
Extracellular matrix.
SIGNAL 1 22
CHAIN 23 422
DOMAIN 23 386
TRANSMEM 387 407
DOMAIN 408 422
POTENTIAL.
ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CTOPLASMIC (POTENTIAL).

Query Match	59.18;	Score 143;	DB 1;	Length 422;
Best Local Similarity	56.18;	Pred. No. 2.5e-10;		
Matches	23;	Conservative	7.	Mismatch

		mismatches	Indels	Gaps
QY	1	5	0	0
	SWPPEVGADICCCCKRGDCGTPSHSRQPPHYMSQMSYS	41		
	: :: : : : :			
Db	307	SWSPVEGDALVCGCCSSGDCSSSRKIQAHGVQWPKRSAS	347	

RESULT	7
ZP3_FELCA	
ID	ZP3_FELCA
AC	Ba0023.
STANDARD:	
PRT;	424 AA

140932, DT
 01-FEB-1996 (Rel. 33, Created)
 01-FEB-1996 (Rel. 33, Last sequence update)
 30-MAY-2000 (Rel. 39, Last annotation update)
 ZONA PELUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELUCIDA GLYCOPROTEIN ZP) (SPERM RECEPTOR) (ZONA PELUCIDA PROTEIN C).
 ZP3 OR ZPC.
 Fetus silvestris catus (cat).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Fels.
 NCBI_TaxID:9685;

(1)
 SEQUENCE FROM N.A.
 TISSUE-Ovary;
 MEDLINE-95143578; PubMed-781460;
 Harris J.D., Hilder D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
 Sacco A.G.;
 Cloning and characterization of zona pellucida genes and cDNAs from
 a variety of mammalian species: the ZPA, ZPB and ZPC gene families.*
 DNA Seq. 4:361-393(1994).

[2]
 SEQUENCE FROM N.A.
 TISSUE-Ovary:
 Okazaki Y., Sugimoto M.;
 Submitted (JAN-1995) to the EMBL/Genbank/DBJ databases.
 -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
 SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
 SPECIES-SPECIFICITY OF THE INSEMINATION (BY SIMILARITY).
 -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
 WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
 MATRIX.
 -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES (BY
 SIMILARITY).
 -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on the use of non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; U05778; AAA74390.1; -;
DR EMBL; D45068; BAA08096.1; -;
DR InterPro; IPR001507; -;
DR Pfam; PF00100; zona pellucida. 1.
DR PRINTS; PR00023; zPELUCIDA.
DR PROSITE; PS00682; ZP_DOMAIN; FALSE_NEG.
KW Glycoprotein; Signal; Sulfatase; Sperm;
KW Extracellular matrix; Multigene family;
KW Receptor; Transmembrane;

```

FE SIGNAL 1 22 POTENTIAL.
FE CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
FE DOMAIN 23 383 EXTRACELLULAR (POTENTIAL).
FE TRASMEM 384 404 POTENTIAL.
FE DOMAIN 405 424 CYTOPLASMIC (POTENTIAL).
FE DOMAIN 43 305 ZP.
FE CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FE CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FE CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FE CONFLICT 72 72 D -> W (IN REF. 2).
FE CONFLICT 264 264 D -> Y (IN REF. 2).
SQ SEQUENCE 424 AA: 46853 MW: CFC62F35F9AFC0D CRC64;

Query Match 56.2%; Score 136; DB 1; Length 424;
Best Local Similarity 55.0%; Pred. No. 1,8e-09;
Matches 22; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 2 WFPVGGPADICCCCKGCGTSPHSRRQPHVMSQMSRSVS 41
DB 307 WFPVGGPADICCCCKGCGTSPHSRRQPHVMSQMSRSVS 346

RESULT 8
ZP3_MOUSE STANDARD; PRT: 424 AA.
ID ZP3_MOUSE
MC P10761.
CT 01-JUL-1989 (Rel. 11, Created)
ET 01-NOV-1997 (Rel. 35, Last sequence update)
ET 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA
DE GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
CN ZP3 OR ZP-3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Liver;
RC MEDLINE=88242926; PubMed=3378665;
RX Ringuette M.J., Chamberlin M.E., Baur A.W., Sobieski D.A., Dean J.;
RX "Molecular analysis of cDNA coding for ZP3, a sperm binding protein
RT of the mouse zona pellucida.";
RN Dev. Biol. 127:287-295(1988).
RN [2]
RP REVISION TO 387.
RA Dean J.;
RN Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Liver;
RC MEDLINE=89240048; PubMed=2541416;
RX Kinloch R.A., Rollier R.J., Fimiani C.M., Wassarman D.A.,
RX Wassarman P.M.;
RT "Primary structure of the mouse sperm receptor polypeptide determined
RT by genomic cloning.";
RN Proc. Natl. Acad. Sci. U.S.A. 85:6409-6413(1988).
RN [5]
RP SEQUENCE OF 49-63; 197-204; 219-233 AND 261-275.
RC STRAIN=CD-1;
RC MEDLINE=93050795; PubMed=1330788;
RX Rosiere T.K., Wassarman P.M.;
RX "Identification of a region of mouse zona pellucida glycoprotein mZP3
RT that possesses sperm receptor activity.";
RN Dev. Biol. 154:309-317(1992).
RN [1]
RP FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR

```

```

CC CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
CC CC SPECIES-SPECIFICITY OF THE INSEMINATION.
CC CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
CC CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXTRACELLULAR
CC CC MATRIX.
CC CC -1- TISSUE SPECIFICITY: OOCYTES.
CC CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING THE 2-WEEK GROWTH PHASE OF
CC CC OOGENESIS, PRIOR TO OVULATION.
CC CC -1- SIMILARITY: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.
CC CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC CC or send an email to license@ebi.ac.uk).
CC CC -----
DR EMBL: M20026; AAB18628.1; -.
DR EMBL: X14376; CAA32550.1; -.
DR PIR: S04189; S04189.
DR PIR: A30334; A30334.
DR PIR: A31232; A31232.
DR MGD: MGI:99215; ZP3.
DR InterPro: IPR001507; -.
DR Pfam: PF00100; zona_pellucida; 1.
DR PRINTS: PR00023; ZPELLUCIDA.
DR PROSITE: PS00682; ZP DOMAIN; 1.
KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
KW Extracellular matrix
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 424 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
FT DOMAIN 23 387 EXTRACELLULAR (POTENTIAL).
FT TRASMEM 388 408 POTENTIAL.
FT DOMAIN 409 424 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 308 ZP.
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 424 AA: 46303 MW: 9089903FBD268365 CRC64;

Query Match 52.1%; Score 126; DB 1; Length 424;
Best Local Similarity 53.7%; Pred. No. 2,9e-08;
Matches 22; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 SWFPVGGPADICCCCKGCGTSPHSRRQPHVMSQMSRSVS 41
DB 309 SWFPVGGPADICCCCKGCGTSPHSRRQPHVMSQMSRSVS 349

RESULT 9
ZP3_PIG STANDARD; PRT: 421 AA.
ID ZP3_PIG
AC P42098;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3-BETA PRECURSOR (ZONA PELLUCIDA
DE GLYCOPROTEIN ZP3-BETA) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
GN ZP3B OR ZPC.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RN TISSUE=ovary;
RC Yurewicz E.C., Hilder D., Fontenot G.K., Harris J.;

```

RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RA Okazaki Y., Sugimoto M.;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
 CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
 CC SPECIES-SPECIFICITY OF THE INSEMINATION (BY SIMILARITY).
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
 CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
 CC MATRIX.
 CC -1- TISSUE SPECIFICITY: OOCYTES.
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L22169; AAA31145.1; -
 CC EMBL: D45065; BAA08093.1; -
 CC GLCOSULTEDB: P42098; -
 CC InterPro: IPR001507; -
 CC Pfam: PF00100; zona_pellucida; 1.
 CC PRINTS: PR00023; ZPELUCIDA.
 CC PROSITE: PS00682; ZP_DOMAIN, 1.
 CC DR GLYCOPROTEIN; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
 CC KW Extracellular matrix.
 CC FT SIGNAL 1 21
 CC FT CHAIN 1 22 421
 CC FT DOMAIN 22 381 POTENTIAL.
 CC FT TRANSMEM 382 402 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3-
 CC FT DOMAIN 403 421 BETA.
 CC FT DOMAIN 403 421 EXTRACELLULAR (POTENTIAL).
 CC FT DOMAIN 403 421 CYTOPLASMIC (POTENTIAL).
 CC FT CARBOHYD 124 124 ZP.
 CC FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CONFLICT 101 101 MISSING (IN REF. 2).
 CC FT CONFLICT 107 107 D -> V (IN REF. 2).
 CC FT CONFLICT 163 164 VF -> A (IN REF. 2).
 CC FT CONFLICT 404 404 P -> A (IN REF. 2).
 CC FT SEQUENCE 421 AA; 46239 MW; DC97D760E985213A CRC64;
 SQ
 Query Match 49.28; Score 119; DB 1; Length 421;
 Best Local Similarity 70.48; Pred. No. 2e-07;
 Matches 19; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 OY 2 WFPVGGPADICQCCNKKGCGTSPSHSR 28
 DB 308 WSPVEGPAVICRCCHKGCGTSPSLSRK 334
 ID ZP3_BOVIN STANDARD; PRT; 421 AA.
 AC P48830;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA
 DE GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C).
 DE ZP3 OR ZPC.
 OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Graniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 CC Bovidae; Bovinae; Bos.
 CC NCBI_TaxID:9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE:95143578; Pubmed-7841460;
 RA Harris J.D., Hblier D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
 RA Sacco A.G.;
 RT "Cloning and characterization of zona pellucida genes and cDNAs from
 RT a variety of mammalian species: the ZPA, ZPB and ZPC gene families.";
 RL DNA Seq. 4:361-393(1994).
 CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
 CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
 CC SPECIES-SPECIFICITY OF THE INSEMINATION (BY SIMILARITY).
 CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
 CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. EXTRACELLULAR
 CC MATRIX.
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U05775; AAA74385.1; -
 CC InterPro: IPR001507; -
 CC Pfam: PF00100; zona_pellucida; 1.
 CC PRINTS: PR00023; ZPELUCIDA.
 CC PROSITE: PS00682; ZP_DOMAIN, 1.
 CC DR GLYCOPROTEIN; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
 CC KW Extracellular matrix; Multigene family.
 CC FT SIGNAL 1 22
 CC FT CHAIN 1 23 421
 CC FT DOMAIN 23 381 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
 CC FT TRANSMEM 382 402 EXTRACELLULAR (POTENTIAL).
 CC FT DOMAIN 403 421 POTENTIAL.
 CC FT DOMAIN 403 421 CYTOPLASMIC (POTENTIAL).
 CC FT CARBOHYD 124 124 ZP.
 CC FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT SEQUENCE 421 AA; 46545 MW; 905C4722B7B11DC CRC64;
 SQ
 Query Match 45.08; Score 109; DB 1; Length 421;
 Best Local Similarity 56.78; Pred. No. 3.2e-06;
 Matches 17; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
 OY 2 WFPVGGPADICQCCNKKGCGTSPSHSRPH 31
 DB 308 WSPVEGPTDICRCCKGRCGIGSRMRLSH 337
 ID ZP3_RABBIT STANDARD; PRT; 415 AA.
 AC P48833;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE ZONA PELLUCIDA SPERM-BINDING PROTEIN 3 PRECURSOR (ZONA PELLUCIDA
 DE GLYCOPROTEIN ZP3) (SPERM RECEPTOR) (ZONA PELLUCIDA PROTEIN C)
 DE (FRAGMENT).
 DE ZP3 OR ZPC.
 OS Oryctolagus cuniculus (Rabbit).

CC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
CC Mammalia: Eutheria: Lagomorpha: Leporidae: Oryctolagus.
CX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95143578; PubMed=7841460;
RA Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
RT "Cloning and characterization of zona pellucida genes and cDNAs from
CC a variety of mammalian species: the ZPA, ZPB and ZPC gene families.";
RL DNA Seq. 4:361-393(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U05782; AAA74392.1; -
DR InterPro; IPR001507; -
DR Pfam; PF00100; zona_pellucida; 1.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Glycoprotein; Signal; Sulfatation; Sperm; Receptor; Transmembrane;
KW Extracellular matrix; Multigene family.
FT NON_TER 1
FT SIGNAL <1 18 POTENTIAL.
FT CHAIN 19 415 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
FT DOMAIN 19 378 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 379 399 POTENTIAL.
FT DOMAIN 400 415 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 41 301 ZP.
SQ SEQUENCE 415 AA; 44987 MW; 77396CF1BA3F5CB CRC64;
Cy 1 SMPVPGPADICCCCKGDC---GTP-----SHSRROPHVMSO 35
ID 302 SMAPEVSADICECCGNGCDLLAGSPMNQNHARSLSRKHVTEE 348
CB
RESULT 12
ESG3_HUMAN STANDARD: PRT: 999 AA.
AC P32926;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DESMOGLEIN 3 PRECURSOR (130 KDA PEMPHIGUS VULGARIS ANTIGEN) (PVA).
GN DSG3.
OS Homo sapiens (Human).
CC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
CC Mammalia: Eutheria: Primates: Catarrhini: Homidae: Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92069753; PubMed=1720352;
RA Anagat M., Klaus-Kovtun V., Stanley J.R.;
RT "Autoantibodies against a novel epithelial cadherin in pemphigus
CC vulgaris, a disease of cell adhesion.";
RL Cell 67:869-877(1991).
CC -1- FUNCTION: COMPONENT OF INTERCELLULAR DESMOsome JUNCTIONS.
CC INVOLVED IN THE INTERACTION OF PLAQUE PROTEINS AND INTERMEDIATE
CC FILAMENTS MEDIATING CELL-CELL ADHESION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: EPIDERMIS, TONGUE, TONSIL, ESOPHAGUS AND
CC CARCINOMAS.
CC -1- DOMAIN: CALCIUM MAY BE BOUND BY THE CADHERIN-LIKE REPEATS

CC (POTENTIAL).
CC -1- DISEASE: PEMPHIGUS VULGARIS (PV) IS A POTENTIALLY LETHAL SKIN
CC DISEASE IN WHICH EPIDERMAL BLISTERS OCCUR AS THE RESULT OF THE
CC LOSS OF CELL-CELL ADHESION CAUSED BY THE ACTION OF AUTOANTIBODIES
CC AGAINST DSG3.
CC -1- SIMILARITY: BELONGS TO THE CADHERIN FAMILY, DESMOsome SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M76482; AAA60230.1; -
DR PIR; A41088; IJHUG3.
DR HSSP; P09803; LEDH.
DR MIM; 169615; -
DR InterPro; IPR002126; -
DR Pfam; PF00028; cadherin; 4.
DR PROSITE; PS00232; CADHERIN_1; 3.
KW Cell adhesion; Signal; Transmembrane; Cytoskeleton; Glycoprotein;
KW Calcium-binding; Repeat.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 49 POTENTIAL.
FT CHAIN 50 999 DESMOGLEIN 3.
FT DOMAIN 50 615 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 616 640 POTENTIAL.
FT DOMAIN 641 999 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 50 158 CADHERIN 1.
FT DOMAIN 159 268 CADHERIN 2.
FT DOMAIN 269 383 CADHERIN 3.
FT DOMAIN 386 499 CADHERIN 4.
FT REPEAT 910 935 DESMOGLEIN REPEAT 1.
FT REPEAT 936 966 DESMOGLEIN REPEAT 2.
FT CARBOHYD 110 110 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAc. . .) (POTENTIAL).
SQ SEQUENCE 999 AA; 107503 MW; 60479DD46AC219A1 CRC64;
Cy Query Match 24.4%; Score 59; DB 1; Length 999;
Best Local Similarity 39.3%; Pred. No. 7;
Matches 11; Conservative 4; Mismatches 7; Indels 6; Gaps 1;
QY 10 DICOCNKGDCGT-----PSHSRROPH 31
ID 585 EVCQDNRGICGTSTPTTSPGRTYRPH 612
DB
RESULT 13
5H6_RAT STANDARD: PRT: 436 AA.
ID P31388;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 5-HYDROXYTRYPTAMINE 6 RECEPTOR (5-HT-6) (SEROTONIN RECEPTOR)
GN HTR6.
OS Rattus norvegicus (Rat).
CC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
CC Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Striatum;
RX Medline=93196608; PubMed=7680751;
RA Monsma F.J. Jr., Shen Y., Ward R.P., Hamblin M.W., Sibley D.R.;
RT "Cloning and expression of a novel serotonin receptor with high
affinity for tricyclic psychotropic drugs.";

RL Mol. Pharmacol. 43:320-327(1993).
 RP [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93277562; PubMed-8389146;
 RA Kuat M., Traillort E., Arrang J.-M., Tardivel-Lacombe J., Diaz J.,
 RA Leurs R., Schwartz J.-C.;
 RT "A novel rat serotonin (5-HT6) receptor: molecular cloning,
 RT localization and stimulation of cAMP accumulation.";
 RL Biochem. Biophys. Res. Commun. 193:268-276(1993).
 RP [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WISTAR;
 RA Martini R.;
 CC Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR
 CC 5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTION
 CC AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT STIMULATES ADENYLATE
 CC CYCLASE. IT HAS A HIGH AFFINITY FOR TRICYCLIC PSYCHOTROPIC DRUGS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: LOCALIZED EXCLUSIVELY IN THE CENTRAL NERVOUS
 CC SYSTEM, PREDOMINANTLY IN THE CORPUS STRIATUM BUT ALSO IN VARIOUS
 CC LIMBIC AND CORTICAL REGIONS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L03202; AAA40618.1; -;
 DR EMBL: S62043; AAB26908.1; -;
 DR EMBL: L19365; AAA40611.1; -;
 DR GCRDB: GCR_0723; -;
 DR GCRDB: GCR_0719; -;
 DR InterPro: IPR000226; -;
 DR InterPro: IPR002232; -;
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PRO0237; GPCRHOOPS.
 DR PRINTS: PRO1102; SH6RECEPT.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 34
 FT TRANSMEM 35 57
 FT DOMAIN 58 64
 FT TRANSMEM 65 85
 FT DOMAIN 86 100
 FT TRANSMEM 101 122
 FT DOMAIN 123 144
 FT TRANSMEM 145 166
 FT DOMAIN 167 184
 FT TRANSMEM 185 208
 FT DOMAIN 209 265
 FT TRANSMEM 266 290
 FT DOMAIN 291 295
 FT TRANSMEM 296 320
 FT DOMAIN 321 436
 FT CARBOHYD 9 9
 FT DISULFID 99 180
 FT CONFLICT 57 57
 FT CONFLICT 336 436
 FT SEQUENCE 436 AA; 46922 MW; 6BC00F6A3CBA55FB8 CMC64;

Query Match 22.9%; Score 55.5; DB 1; Length 436;
 Best Local Similarity 34.9%; Pred. No. 9.4;
 Matches 15; Conservative 3; Mismatches 16; Indels 9; Gaps 2;
 Oy 4 PVOGPADITQ-----CCNKGDGCTPSHRSRQPRVMSQMSNSVS 41
 Db 382 PLOGAPRACSSQPSFCCLEPRPPTPRHPGPP---LWSTSL 420
 RESULT 14
 ID PML_HUMAN STANDARD; PRT; 865 AA.
 AC 043490;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROMININ-LIKE PROTEIN 1 PRECURSOR (ANTIGEN AC133).
 GN PROM1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-42; 200-211; 280-291 & 641-656.
 RC TISSUE=Fetal liver, AND Retinoblastoma;
 RX MEDLINE-98052559; PubMed-9389721;
 RA Miraglia S., Godfrey W., Yin A.H., Atkins K., Warnke R., Holden J.T.,
 RA Bray R.A., Waller E.K., Buck D.W.;
 RT "A novel five-transmembrane hematopoietic stem cell antigen:
 RT isolation, characterization, and molecular cloning.";
 RL Blood 90:5013-5021(1997).
 RN [2]
 RP CHARACTERIZATION.
 RC TISSUE=Fetal liver;
 RX MEDLINE-98052558; PubMed-9389720;
 RA Yin A.H., Miraglia S., Zanjanj E.D., Almeida-Porada G., Ogawa M.,
 RA Leary A.G., Olweus J., Kearney J., Buck D.W.,
 RA "AC133, a novel marker for human hematopoietic stem and progenitor
 RT cells.";
 RL Blood 90:5002-5012(1997).
 RN [3]
 RP DISFASE.
 RX MEDLINE-20056119; PubMed-10587575;
 RA Maw M.A., Corbelle D., Koch J., Hellwig A., Wilson-Wheeler J.C.,
 RA Bridges R.J., Kumarmanickavel G., John S., Nancarrow D., Roeper K.,
 RA Weigmann A., Hutterer W.B., Denton M.J.;
 RT "A frameshift mutation in prominin (mouse)-like 1 causes human retinal
 RT degeneration.";
 RL Hum. Mol. Genet. 9:27-34(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED ON CD34 HEMATOPOIETIC
 CC STEM AND PROGENITOR CELLS IN ADULT AND FETAL BONE MARROW, FETAL
 CC LIVER, CORD BLOOD AND ADULT PERIPHERAL BLOOD. NOT DETECTED ON
 CC OTHER BLOOD CELLS. ALSO EXPRESSED IN A NUMBER OF NONLYMPHOID
 CC TISSUES INCLUDING RETINA, PANCREAS, PLACENTA, KIDNEY, LIVER, LUNG,
 CC BRAIN AND HEART.
 CC -1- PTM: GLYCOSYLATED.
 CC -1- DISEASE: DEFECTS IN PROM1 ARE THE CAUSE OF AN AUTOSOMAL RECESSIVE
 CC FORM OF RETINAL DEGENERATION CHARACTERIZED BY NIGHT BLINDNESS AND
 CC LOSS OF PERIPHERAL VISION FROM CHILDHOOD WITH PROGRESSION TO
 CC PROFOUND VISUAL IMPAIRMENT AND EXTINGUISHED ELECTRORETINOGRAMS BY
 CC THE THIRD DECADE. A SINGLE NUCLEOTIDE DELETION YIELDS A TRUNCATED
 CC PROTEIN LACKING HALF OF THE SECOND EXTRACELLULAR LOOP, THE FINAL
 CC MEMBRANE-SPANNING DOMAIN AND THE CYTOPLASMIC C-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE PROMININ FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

or send an email to license@slsb.sib.ch).
-----
Cc      EMBL; AF027208; AAB92514.1; -
Dr      MIM: 604365; -
KW      Transmembrane; glycoprotein; Antigen; signal; polymorphism;
        vision.
FT      SIGNAL                     1          19          POTENTIAL.
FT      CHAIN                     20          108          PROMININ-LIKE PROTEIN 1.
FT      DOMAIN                    109          129          EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM                    130          157          POTENTIAL.
FT      DOMAIN                    158          178          CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM                    179          433          POTENTIAL.
FT      DOMAIN                    434          454          EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM                    455          486          POTENTIAL.
FT      DOMAIN                    487          507          CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM                    508          792          POTENTIAL.
FT      TRANSMEM                    793          813          EXTRACELLULAR (POTENTIAL).
FT      DOMAIN                    814          865          POTENTIAL.
FT      CARBOHYD                    220          220          CYTOPLASMIC (POTENTIAL).
FT      CARBOHYD                    274          274          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD                    395          395          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD                    414          414          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD                    548          548          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD                    580          580          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD                    729          729          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD                    730          730          N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      VARIANT                    31          31          A -> G.
FT      VARIANT                    31          31          /FTid-VAR_010382.
FT      VARIANT                    31          31          A -> S.
FT      VARIANT                    31          31          /FTid-VAR_010383.
FT      CONFLICT                    200          200          D -> P OR A (IN AA SEQUENCE).
FT      CONFLICT                    284          284          S -> R (IN AA SEQUENCE).
FT      CONFLICT                    288          288          S -> D (IN AA SEQUENCE).
SC      SEQUENCE                    865 AA; 97201 MM; D21C8C05ADB2DEDF CRC64;

Query Match          22.7%; Score 55; DB 1; Length 865;
Best Local Similarity 47.4%; Pred. No. 19;
Matches 9; Conservative 4; Mismatches 4; Indels 2; Gaps 2;

OY      11 ICOCCKGDCGPHSHSRRO 29
Db      132 WCRCCK--CGGEMHOROK 148

RESULT 15
PROM_MOUSE
ID      PROM_MOUSE          STANDARD:          PRT:          867 AA.
AC      054990; 035408;
DT      01-OCT-2000 (Rel. 40, Created)
DT      01-OCT-2000 (Rel. 40, Last sequence update)
DT      01-OCT-2000 (Rel. 40, Last annotation update)
DE      PROMININ PRECURSOR (ANTIGEN AC133 HOMOLOG).
EN      PROM.
GN      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCI_TaxID=10090;
RX      NCI_TaxID=10090;
RP      SEQUENCE FROM N.A.
RC      TISSUE-Brain;
RC      TISSUE-Brain;
RL      Miraglia S., Godfrey W., Buck D.;
RN      Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
RP      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE-Kidney;
RC      MEDLINE=98024147; PubMed=9356465;
RX      Weigmann A., Corbail D., Helliwig A., Huttner W.B.;
RA      "Prominin, a novel microvilli-specific polytopic membrane protein of
RT      the apical surface of epithelial cells, is targeted to plasmalemmal
RL      protrusions of non-epithelial cells."
Proc. Natl. Acad. Sci. U.S.A. 94:12425-12430(1997).

```

CC	-1-	SUBCELLULAR LOCATION:	INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC	-1-	TISSUE SPECIFICITY:	IN THE EMBRYO, EXPRESSED ON THE APICAL SIDE OF NEUROEPITHELIAL CELLS AND OF OTHER EPITHELIA SUCH AS LUNG BUDS, GUT AND URETERH BUDS. IN THE ADULT, EXPRESSED AT THE APICAL SIDE OF THE KIDNEY TUBULES AND OF THE EPENDYMAL LAYER OF THE BRAIN. NOT EXPRESSED IN GUT, LIVER, LUNG, PITUITARY, ADRENAL, HEART OR SPLEEN.
CC	-1-	SIMILARITY:	BELONGS TO THE PROMININ FAMILY.
CC			-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outpost - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb.sib.ch).	
CC		EMBL; AF036633; AAB86916.1; -	
DR		EMBL; AF026269; AAB86715.1; -	
DR	MGI:	1100886;	Prom.
KM	Signal:	Transmembrane; Glycoprotein.	
FT	SIGNAL	1	POTENTIAL.
FT	CHAIN	20	PROMININ.
FT	DOMAIN	20 107	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	108 128	POTENTIAL.
FT	DOMAIN	129 158	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	159 179	POTENTIAL.
FT	DOMAIN	180 434	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	435 455	POTENTIAL.
FT	DOMAIN	456 487	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	488 508	POTENTIAL.
FT	DOMAIN	509 794	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	795 815	POTENTIAL.
FT	DOMAIN	816 867	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	273 273	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	291 291	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	332 332	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	374 374	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	415 415	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	554 554	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	581 581	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	732 732	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	64	S -> N (IN REF. 2).
FT	CONFLICT	84	K -> N (IN REF. 2).
FT	CONFLICT	94	K -> L (IN REF. 2).
FT	CONFLICT	668 668	MISSING (IN REF. 2).
FT	CONFLICT	844 844	G -> D (IN REF. 2).
EQ	SEQUENCE	867 AA; 97112 MW; D4426372552B3C8 CRC64;	

Search completed: July 12, 2001, 08:27:09
Job time: 179 sec

```

Query Match          22.7%; Score 55; DB 1; Length 867;
Best Local Similarity 47.4%; Pred. No. 19;
Matches 9; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

QY 11 ICOCCKGDCCTPSHSRQ 29
      :|:|:|:| 11 | 1:
DB 133 MCRCCK--CGGEMHQRQK 149

Search completed: July 12, 2001, 08:27:09
Job time: 179 sec

```


1	211	89.7	210	4	012903
2	125	51.7	424	11	p97708
3	115	47.5	418	11	Q9ER10
4	100.5	41.5	454	13	Q91675
5	100.5	41.5	460	13	Q91728
6	98.5	40.7	446	13	P79762
7	98.5	40.7	446	13	Q9PWF8
8	95	39.3	422	6	077685
9	87	36.0	445	13	073670
10	60.5	25.0	99	2	Q9PE89
11	60.5	25.0	436	13	Q9DG37
12	59.5	24.6	934	2	005478
13	57	23.6	1120	5	Q20778
14	56.5	23.3	819	2	Q9PK14
15	56	23.1	122	4	Q9P166
16	56	23.1	448	10	Q9FVY3
17	56	23.1	641	11	008463
18	56	23.1	642	11	070421
19	56	23.1	647	4	Q94815
					094815 homo sapien
					Q12903 homo sapien
					p97708 rattus norv
					Q9ER10 microtus dr
					Q91675 xenopus lae
					Q91728 xenopus lae
					P79762 gallus galli
					Q9PWF8 gallus galli
					077685 trichosurus
					073670 coturnix co
					Q9PE89 xyliella fas
					Q9DG37 plimethales
					005478 streptomyc
					Q20778 caenorhabd
					Q9PK14 chlamydia m
					Q9P166 homo sapien
					Q9FVY3 saccharum c
					008463 rattus norv
					070421 mus musculu
					094815 homo sapien

20	56	23.1	648	4	09UP38	09up38 homo sapien
21	55.5	22.9	344	11	054788	054788 mus musculus
22	55.5	22.9	435	13	09PU36	09puj6 carassius a
23	55.5	22.9	447	10	09ZRP9	09zrp9 malus domes
24	55	22.7	295	2	044588	044588 alcaligenes
25	55	22.5	591	11	09J149	09j149 rattus norv
26	54.5	22.7	340	14	09JF36	09jf36 vaccinia vir
27	54.5	22.7	424	13	09Z027	09z027 cyprinus car
28	54.5	22.5	559	13	0919M5	0919m5 xenopus lae
29	54.5	22.5	735	11	060473	060473 cavia porcea
30	54	22.3	214	2	09RD56	09rd56 streptomyce
31	54	22.3	214	6	097602	097602 oryctolagus
32	54	22.3	315	6	097601	097601 oryctolagus
33	54	22.3	477	6	097600	097600 oryctolagus
34	54	22.3	537	4	09HBE2	09hbe2 homo sapien
35	54	22.3	544	1	048304	048304 halofexa s
36	54	22.3	592	13	057328	057328 gallus gall
37	54	22.3	609	4	09HDE1	09hde1 homo sapien
38	53.5	22.1	227	5	09V4G7	09v4g7 drosophila
39	53.5	22.1	351	4	060521	060521 homo sapien
40	53.5	22.1	338	1	09YA58	09ya58 aemopyrum p
41	53.5	22.1	379	10	09FYF9	09fyf9 arabidopsis
42	53.5	22.1	738	4	060357	060357 homo sapien
43	53.5	22.1	2168	5	019350	019350 caenorhabdl
44	53	21.9	197	6	028554	028554 ovis aries
45	53	21.9	226	10	09SBR2	09sbr2 fragaria an

RESULT	1	
ID	012903	PRELIMINARY; PRT: 210 AA.
AC	Q12903;	
DT	01-NOV-1996 (TREMBLrel. 01, Created)	
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)	
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)	
DE	POM-zp3.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=OVARY;	
RX	MEDLINE=95309900; PubMed=7789967;	
RA	Kipersztok S., Osawa G.A., Liang L.F., Modi W.S., Dean J.;	
RT	"POM-zp3", a bipartite transcript derived from human zp3 and a POM121	
RT	homologue.";	
RL	Genomics 25:354-359(1995).	
DR	EMBL: U10099; AAA85788.1; -	
DR	InterPro: IPR001507; -	
DR	Pfam: PF00100; zona_pellucida.1	
QO	SEQUENCE 210 AA; 23196 MW; 663723ECLB77C8D0 CRC64;	

Query Match	89.7%;	Score 217;	DB 4;	Length 210;
Best Local Similarity	87.8%;	Pred. No.	3.6e-22;	
Matches	36;	Conservative	2;	Mismatches 3;
				Indels 0; Gaps 0;
QY	1	SWPVGADPICCCCKNGDCGPHSHRRQPHVMSQWSRVS	41	
Db	146	SWPEVEGPADICCCCKNGDCGPHSHRRQPRVYSQWSTPAS	186	
RESULT	2			
P97708				
ID	P97708	PRELIMINARY;	PRT;	424 AA.
AC	P97708;	05084;		
DT	01-NOV-1998	(TREMBLrel. 08,	created)	
DT	01-NOV-1998	(TREMBLrel. 08,	Last sequence update)	
DT	01-MAR-2001	(TREMBLrel. 16,	Last annotation update)	

OS	Microtus brandti.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Arvicolinae;
OC	Microtus.
OX	NCBI_TaxID=137787;
RN	(1)
RP	SEQUENCE FROM N.A.
RC	TISSUE=OVARY;
RA	Lf H., Piao Y.S., Zhang Z.B., Zhu C., Hardy C., Hinds L.A.:
RT	"Cloning and characterization of zona pellucida cDNA from Brandt's
RL	vole";
DR	Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
SO	SEQUENCE 418 AA; 45662 MW; 12C0506255F46BF CRC64;
QY	1 SWFPVQGPDICGCCNKGDGCTPSRSRRQPHVMQSWMR 38
Db	309 SWLPVEGDTDVCDCCTKGCSSSRYSRPRAHAVAPRNMR 346
<pre> 1:1 11 1111 : 11 : 1:: 1 </pre>	
Matches 17; Conservative 8; Mismatches 13; Indels 0; Gaps 0;	
RESULT 4	
ID 091675	PRELIMINARY; PRT; 454 AA.
AC 091675;	
DT 01-NOV-1996 (TREMBLrel. 01, Created)	
FT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)	
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
DE ZONA PELLUCIDA C GLYCOPROTEIN PRECURSOR.	
GN XLZPC.	
OS Xenopus laevis (African clawed frog).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;	
OC Xenopodinae; Xenopus.	
OX NCBI_TaxID=8355;	
RN RP	[1]
RP SEQUENCE FROM N.A.	
RL Hedrick J.L., Yang J.C.;	
RL Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.	
DR EMBL; U44952; AAB39079.1; -.	
DR InterPro; IPR001507; -.	
DR Pfam; PF00100; zona_pellucida; 1.	
DR SMART; SM00241; ZP; 1.	
KW SIGNAL.	
FT CHAIN	1 21 POTENTIAL.
FT CHAIN	22 454 ZONA PELLUCIDA C GLYCOPROTEIN.
SO SEQUENCE 454 AA; 49589 MW; A8DB9A5B978E5197 CRC64;	
Query Match	41.5%; Score 100.5; DB 13; Length 454;
Best Local Similarity	46.3%; Pred. No. 6.2e-06;
Matches 19; Conservative 9; Mismatches 10; Indels 3; Gaps 3;	
QY	1 SWFPVQGPDICGCCNKGDG-CGPSHSRR-OPIVN-SQMSR 38
Db	330 SMSPLQGSNTSCCDTGNCSVPQSGSRLRGPFYFGSMNQ 370
<pre> : : :1:1 1 111 1: 1:1: </pre>	
Matches 19; Conservative 9; Mismatches 10; Indels 3; Gaps 3;	
RESULT 5	
ID 091728	PRELIMINARY; PRT; 460 AA.
AC 091728;	
DT 01-NOV-1996 (TREMBLrel. 01, Created)	
FT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)	
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
DE GP43.	
OS Xenopus laevis (African clawed frog).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;	

```
OC Xenopodinae; Xenopus.
OK NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-OVARIES FROM HCG-TREATED FEMALES;
RX MEDLINE=98013448; PubMed=9352194;
RA Kubo H., Kawano T., Tsubuki S., Kawashima S., Katagiri C., Suzuki A.;
RT "A major glycoprotein of Xenopus egg vitelline envelope, gp41, is a
RL frog homolog of mammalian ZP3.";
DEV. Growth Differ. 39:405-417(1997).
DR EMBL: D86568; BAA13117.1; -.
DR InterPro: IPR001507; -.
DR Pfam: PF00100; zona_pellucida; 1.
D3 SMART: SM00241; ZP; 1.
SQ SEQUENCE 460 AA; 50084 MW; 3798F3E3394FDF34 CRC64;

Query Match 41.5%; Score 100.5; DB 13; Length 460;
Best Local Similarity 46.3%; Pred. No. 6.3e-06;
Matches 19; Conservative 9; Mismatches 10; Indels 3; Gaps 3;

QY 1 SWFVQGPADICCCCKNGDC-GTPSHSR-OPHYM-SQMSR 38
D3 336 SWSPLOGPSNICSCDPTGNCVSPGSRRLGPFSGSRMNO 376
P79762 PRELIMINARY; PRT; 446 AA.
AC P79762;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE ZPC.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OK NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99203469; PubMed=10103002;
RA Takeuchi Y., Nishimura K., Aoki N., Adachi T., Sato C., Kitajima K.,
RA Matsuda T.;
RT "A 42-kDa glycoprotein from chicken egg-envelope, an avian homolog of
the ZPC family glycoproteins in mammalian zona pellucida. Its first
identification, cDNA cloning and granulosa cell-specific expression.";
RL Eur. J. Biochem. 260:736-742(1999).
DR EMBL: D89097; BAA13760.2; -.
DR InterPro: IPR001179; -.
DR Pfam: PF00100; zona_pellucida; 1.
DR PRINTS: PR00023; ZPELUCIDA.
DR PROSITE: PS00453; FKBP_PPIASE_1; UNKNOWN_1.
SQ SEQUENCE 446 AA; 47700 MW; DC541532C22F03C CRC64;

Query Match 40.7%; Score 98.5; DB 13; Length 446;
Best Local Similarity 44.4%; Pred. No. 1.1e-05;
Matches 16; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

QY 1 SWFVQGPADICCCCKNGDCGTPSHSRQPHYMSQW 36
D3 330 TWVPEGSRDVNCCEGTGNCCEPPALSRRL-NPEMRW 364
C9PWF8 PRELIMINARY; PRT; 446 AA.
AC C9PWF8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
```

```
DE ZONA PELLUCIDA C PROTEIN.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OK NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Kono Y., Matsuda T.;
RT "Gallus gallus zona pellucida C protein gene.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB031033; BAA83418.1; -.
DR InterPro: IPR001179; -.
DR InterPro: IPR001507; -.
DR Pfam: PF00100; zona_pellucida; 1.
DR PRINTS: PR00023; ZPELUCIDA.
DR PROSITE: PS00453; FKBP_PPIASE_1; UNKNOWN_1.
DR SMART: SM00241; ZP; 1.
SQ SEQUENCE 446 AA; 47670 MW; DC5409532C22F03C CRC64;

Query Match 40.7%; Score 98.5; DB 13; Length 446;
Best Local Similarity 44.4%; Pred. No. 1.1e-05;
Matches 16; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

QY 1 SWFVQGPADICCCCKNGDCGTPSHSRQPHYMSQW 36
D3 330 TWVPEGSRDVNCCEGTGNCCEPPALSRRL-NPEMRW 364
O77685 PRELIMINARY; PRT; 422 AA.
AC O77685;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ZONA PELLUCIDA 3 PROTEIN.
GN ZP3.
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Diprotodontia; Phalangeridae; Trichosurus.
OK NCBI_TaxID=9337;
RN [1]
RP SEQUENCE FROM N.A.
RA McCartney C.A., Mate K.E.;
RT "Cloning and characterization of a zona pellucida 3 cDNA from a
RT marsupial, the brushtail possum Trichosurus vulpecula.";
RL Zygote 0:0-0(1998).
DR EMBL: AF079524; AAC28736.1; -.
DR InterPro: IPR001179; -.
DR InterPro: IPR001507; -.
DR Pfam: PF00100; zona_pellucida; 1.
DR PRINTS: PR00023; ZPELUCIDA.
DR PROSITE: PS00453; FKBP_PPIASE_1; UNKNOWN_1.
DR SMART: SM00241; ZP; 1.
SQ SEQUENCE 422 AA; 45714 MW; 58BA61A92F82612A CRC64;

Query Match 39.3%; Score 95; DB 6; Length 422;
Best Local Similarity 53.6%; Pred. No. 3.3e-05;
Matches 15; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 WFPVQGPADICCCCKNGDCGTPSHSRQ 29
D3 314 WLPVQGPADICCCCTGTCISLSSSRK 341
O73670 PRELIMINARY; PRT; 446 AA.
AC O73670;
```

DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE OVALI ZPC.
 OS Coturnix coturnix japonica (Japanese quail).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Coturnix.
 NCBI_TaxID=93934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RL Kono Y., Tsuda T., Aoki N., Kitajima K., Matsuda T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB012606; BAA25637.1; -
 DR InterPro; IPR001179; -
 DR InterPro; IPR001507; -
 DR Pfam; PF00100; zona_peptidica; 1.
 DR PRINTS; PR00023; ZPEPTIDICA.
 DR PROSITE; PS00453; FKBP_PPIASE_1; UNKNOWN_1.
 DR SMART; SM00241; ZP; 1.
 SO SEQUENCE 446 AA; 47621 MW; 67A98797A58C135B CRC64;

Query Match 36.0%; Score 87; DB 13; Length 446;
 Best Local Similarity 37.8%; Pred. No. 0.00043;
 Matches 14; Conservative 6; Mismatches 13; Indels 4; Gaps 1;

QY 1 SWFVPGADICCCNKGDCGTPSHSRQPHVMSQWS 37
 DB 330 TWVPEGSRDVCSCEFGNCDAARLR-----LNNWS 362

RESULT 10
 Q9PE89 PRELIMINARY; PRT; 99 AA.
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL PROTEIN XF1139.
 GN XF1139.
 OS Xylella fastidiosa.
 CC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C.
 MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facinon A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnesiel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kurrae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marinho C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matukuma A.Y.,
 RA Menck C.F.M., Miranda E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi W.F., Trufi D., Tsai S.M., Tsunako M.H.,

RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*."
 RL Nature 406:151-159(2000).
 DR EMBL; AE003949; AAF83949.1; -
 KW Hypothetical protein.
 SO SEQUENCE 99 AA; 11303 MW; 2CF720F47433E81 CRC64;

Query Match 25.0%; Score 60.5; DB 2; Length 99;
 Best Local Similarity 45.8%; Pred. No. 0.49;
 Matches 11; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 12 CQCNKGDCGTPSHSRQPHVMSQ 35
 DB 36 CRCARK-HCGLPEHSRTFPIPEQ 58

RESULT 11
 Q9DG37 PRELIMINARY; PRT; 436 AA.
 AC Q9DG37;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE ZONA PEPTIDICA GLYCOPROTEIN 3 (FRAGMENT).
 GN ZP3.
 OS Pimphales promelas.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 CC Cypriniformes; Cyprinidae; Leuciscinae; Pimphales.
 NCBI_TaxID=90988;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Morris C.A., Thomas-Jones R.E., Cryer J., Woodhead S., Kille P.;
 RT "Identification of the ZP3 cDNA sequence from *Pimphales promelas*
 RT exposed to diethylstilbestrol."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF192407; AAG28398.1; -
 FT NON_TER 1
 SO SEQUENCE 436 AA; 47690 MW; 8C54DC948DBC41B6 CRC64;

Query Match 25.0%; Score 60.5; DB 13; Length 436;
 Best Local Similarity 27.5%; Pred. No. 1.8;
 Matches 11; Conservative 4; Mismatches 24; Indels 1; Gaps 1;

QY 2 WFPVPGADICCCNKGDCGTPSHSRQPHVMSQWSRSVS 41
 DB 385 WFAADGNDVCGGCCD-STGCPDGEFASPYGIEWECKAS 423

RESULT 12
 ID 005478 PRELIMINARY; PRT; 934 AA.
 AC 005478;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE LIPA TRANSCRIPTIONAL ACTIVATOR.
 GN LIPR.
 OS Streptomyces sp.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=1931;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ML1.
 MEDLINE=93138418; PubMed=8422993;
 RA Perez C., Juarez K., Garcia-Castells E., Soberon G.,
 RA Servin-Sanzalez L.;
 RT "Cloning, characterization, and expression in *Streptomyces lividans* 66
 RT of an extracellular lipase-encoding gene from *Streptomyces* sp. ML1.";

R EMBL: AF116668; AAF71088.1; .
Q SEQUENCE 122 AA; 13695 MW; B428ED1274CE81CD CRC64;

Query Match 23.1%; Score 56; DB 4; Length 122;
Best Local Similarity 36.6%; Pred. No. 2.4;
Matches 15; Conservative 6; Mismatches 8; Indels 12; Gaps 2;

Y 6 OGPADI-----COCCKNGDCGTPSHSRROPHVMSQMSRSVS 41
:||||| |
b 32 RGPADSSHAFLQC-----FPCHCRPPWPVISRKGRAIN 65

Search completed: July 12, 2001, 08:26:53
Job time: 163 sec